



UNIVERSITY OF GOTHENBURG

The Structural Modularity and Inherent Dynamics of the DegP Protease together with its Intertwined Role in the Bacterial Periplasm

Darius Šulskis

Department of Chemistry and Molecular Biology

Thesis for the degree of Doctor of Philosophy in the Natural Sciences

Gothenburg, 2021

Thesis for the Degree of Doctoral Philosophy in the Natural Sciences

The Structural Modularity and Inherent Dynamics of the DegP Protease together with its
Intertwined Role in the Bacterial Periplasm

Darius Šulskis

Cover: Representation of the *E. coli* periplasm with periplasmic protein structures (PDB ids:
6LYS, 1SG2, 1KY9, 3OU0, 5NG5, 2PV3, 1EQ7)

Copyright ©2021 by Darius Šulskis

ISBN: 978-91-8009-576-1 (PRINT)

ISBN: 978-91-8009-577-8 (PDF)

Available online at <http://hdl.handle.net/2077/69931>

Department of Chemistry and Molecular Biology

Division of Biochemistry and Structural Biology

University of Gothenburg

SE-405 30 Göteborg, Sweden

Borås, Sweden 2021

Printed by Steam Specialtryck AB

The most important step a man can take. It's not the first one, is it?

It's the next one. Always the next step.

Brandon Sanderson, Oathbringer

In Memory of My Father, Rimantas Šulskis

Abstract

The protein quality control machinery is a delicate and integrated network of molecular tools working together to fold or remove unwanted proteins from the cell. A distinct set of “gatekeepers” are involved in this process including molecular chaperones, proteases, oxidoreductases, transferases, and many others in both eukaryotic and prokaryotic organisms. Whereas eukaryotes use a more advanced and multi-layered protein machinery, prokaryotes are more adapted to respond to sudden stresses *via* straight, simple pathways. As a member of gram-negative bacteria, *E. coli* has evolutionarily adapted to have a rich periplasm that protects it against external and internal dangers.

One of the periplasmic “gatekeepers” is the homo-oligomeric DegP-protease that plays a crucial role in the biogenesis and degradation of β -barrel outer-membrane proteins within the periplasmic space. The DegP protein consists of a protease and two regulatory PDZ domains, that control and modulate DegPs association from a 300 kDa hexamer to a 1 MDa cage complex. DegP is activated under heat shock conditions and is well characterized on both the genetic as well as the biochemical level, however, its structural transitions, loop configurations that govern active site regulation as well as dynamical details, remain poorly understood. The aim of my thesis was to delineate the origin of molecular changes of DegP at the atomic level along with DegP's role in the periplasm by using biochemistry and advanced solution NMR techniques.

In this thesis, I characterized an induced temperature switch of DegP that is controlled by transmission from hexamer to trimer at elevated temperatures *via* a stabilizing methionine-aromatic motif in the regulatory PDZ domains. Furthermore, fine-tuned dynamics of the protease domain exposed inherent relaxation of allosteric residues within the protease core as well as the extent of inhibitory LA loop motions. Along assigning for the first time DegP individual domains by solution by NMR spectroscopy approaches, I started out to study full-length DegP by solid-state NMR. Finally, we characterized novel interactions between periplasmic chaperone Skp and Deg proteases. We explicitly observed the degradation of the periplasmic chaperone Skp monomer by both DegP and its homologue DegQ in *in vitro* experiments, revealing potentially a novel layer of regulation in *E. coli* protein quality control. Altogether, I managed to overcome the challenging DegP size for NMR spectroscopy and described its structural and dynamic properties to a level of detail previously not possible.

Sammanfattning på Svenska

Kontrollen av proteinkvalitet i en cell är ett integrerat nätverk av molekyllära verktyg som tillsammans verkar för att klyva och ta bort oönskade proteiner från cellen om dem ej går att reparera. En distinkt uppsättning så kallade "portvakter" är involverade i denna process som inkluderar molekyllära chaperoner, proteaser, oxidoreduktaser, transferaser och många andra maskinerier i både eukaryota och prokaryota organismer. Medan eukaryoter använder ett mer avancerat och flerskiktigt proteinmaskineri, så är prokaryoter mer anpassade att svara på plötsliga påfrestningar via enklare och mindre komplexa vägar. Som en medlem av familjen gramnegativa bakterier har *E. coli* evolutionärt anpassat sig för att ha en välförsedd periplasma som skyddar cellen mot yttre och inre faror.

En av de periplasmatiske "portvakterna" är det homo-oligomera DegP-proteaset som spelar en avgörande roll i biogenes och nedbrytning av yttermembranproteiner bestående av β -fat i det periplasmatiske utrymmet. DegP-proteinet består av en proteas domän och två regulatoriska PDZ-domäner, som kontrollerar och modulerar DegP:s sammansättningen från en 300 kDa hexamer till ett 1 MDa-multimerkomplex. DegP aktiveras under temperaturhöjning och är väl karaktäriserat på både den genetiska såväl som den biokemiska planet. Dess strukturella övergångar och loop-konfigurationer som reglerar den aktiva konformeringen såväl som dynamiska detaljer är fortfarande underförstådda. Syftet med min doktorandtjänst var att definiera ursprunget till molekyllära förändringar av DegP på atomnivå samtidigt få en förståelse av DegP:s roll i periplasman genom att använda biokemiska metoder och framförallt avancerade NMR-tekniker.

I denna avhandling karakteriserade jag en inducerad temperaturväxling av DegP som styrs av överföring från hexamer till trimer vid förhöjda temperaturer via ett stabiliserat metionin-aromatiskt motiv i de regulatoriska PDZ-domänerna. Vidare visade finjusterad dynamik i proteasdomänen inhyst spin-relaxation av allosteriska residuer i proteaskärnan såväl som omfattande hämning av LA-looprörelser. Medan tilldelningen av residuer för individuella DegP-domäner utfördes för första gången med NMR-spektroskopi i lösning lyckades jag också börja studera fullängdsstrukturen av DegP med "solid-state NMR". Slutligen karakteriserade vi nya interaktioner mellan den periplasmatiske chaperonet Skp och Deg-proteaser. Jag observerade tydlig nedbrytning av den periplasmatiske chaperonet Skp:s monomer av både DegP och dess homolog DegQ i *in vitro*-experiment, vilket potentiellt avslöjade en förut okänd

mekanism av reglering i *E. coli*-proteinkvalitetskontroll. Sammanfattningsvis lyckades jag bemästra den utmanande storleken av DegP för NMR-spektroskopi och beskrev dess strukturella och dynamiska egenskaper till en detaljnivå som tidigare inte varit möjligt.

Publications

The thesis consists of the following research papers:

1. Šulskis, D., Thoma, J., Burmann, B.M. 2021. Structural basis of DegP-protease temperature-dependent activation. *Science Advances*, **7**, eaabj1816 (2021). **(Paper I)**
2. Šulskis, D., Burmann, B.M. Functional regulation of the DegP-proteolytic domain by the interplay of dynamic loops. *Manuscript in preparation*. **(Paper II)**
3. Šulskis, D., Vallet, A., Schanda, P., Burmann, B.M. Sequence specific backbone resonance assignments of *E. coli* serine protease DegP by solid-state NMR. *Manuscript in preparation*. **(Paper III)**
4. Šulskis, D., Thoma, J., Burmann, B.M. Proteolytic regulation of Skp chaperone function forms the basis of its interplay with the serine proteases DegQ and DegP. *Manuscript in preparation*. **(Paper IV)**
5. Burmann, B.M., Gerez, J.A., Matečko-Burmann, I., Campioni, S., Kumari, P., Ghosh, D., Mazur, A., Aspholm, E.E., Šulskis, D., Wawrzyniuk, M., Bock, T., Schmidt, A., Rüdiger, S.G.D., Riek, R. and Hiller, S. 2020. Regulation of α -synuclein by chaperones in mammalian cells. *Nature*, **577**, 127–132 (2020). **(Paper V)**

Related papers:

1. Hiller, S., Šulskis, D., Mazur, A., Kawale A.A., Burmann, B.M. GLIMPSE: Global Lifetime Measurements by PRE spin-label exchange. *Manuscript in preparation*. **(Paper VI)**
2. Horvath, I., Blockhuys, S., Šulskis, D., Holgersson, S., Kumar, R., Burmann, B.M. and Wittung-Stafshede, P., 2019. Interaction between copper chaperone Atox1 and Parkinson's disease protein α -Synuclein includes metal-binding sites and occurs in living cells. *ACS Chemical Neuroscience*, **10**, 4659–4668 (2019). **(Paper VII)**

Contribution report

Paper I: I, together with my supervisor and a post-doctoral fellow, planned and designed the research. I performed all the protein productions and biochemical assays. I, together with my two co-authors, analyzed the data and did the sequence-specific assignments. I, together with my two co-authors, prepared and wrote the manuscript.

Paper II: I, together with my supervisor, planned and designed the research. I performed all the protein productions and biochemical assays. I, together with my supervisor, analyzed the data and performed the sequence-specific resonance assignments. I, together with my supervisor, prepared and wrote the manuscript.

Paper III: I, together with my supervisor, planned and designed the research. I performed all the protein productions and biochemical assays. I, together with my supervisor, analyzed the data and performed the sequence-specific resonance assignments. I, together with my supervisor, prepared and wrote the manuscript.

Paper IV: I, together with my supervisor and a post-doctoral fellow, planned and designed the research. I performed all the protein productions. I, together with my supervisor, analyzed the data and did the sequence-specific resonance assignments. I, together with my two co-authors, prepared and wrote the manuscript.

Paper V: I, together with different co-authors of the paper, did the protein production and prepared samples for NMR experiments.

Abbreviations

A.U. –Arbitrary Unit

ATP – Adenosine Triphosphate

Cpx – Conjugative plasmid expression

CSA – Chemical Shift Anisotropy

CYANA – Combined Assignment and Dynamics Algorithm for NMR Applications

DD – Dipole-Dipole

DegP – Periplasmic serine endoprotease

Dsb – Disulphide bond (protein family)

E. coli – *Escherichia coli*

EDTA – Ethylenediaminetetraacetic Acid

FLYA – Fully Automated NMR structure Determination Algorithm

IGFBP – Insulin-like Growth Factor Binding Protein

IM – Inner Membrane

IMAC – Immobilized Metal Affinity Chromatography

IPTG – Isopropyl β - d-1-Thiogalactopyranoside

Gdn-HCl – Guanidinium chloride

HSQC – Heteronuclear Single Quantum Coherence

HtrA – High temperature requirement A

MF – Model-Free

MTSL – (*S*-(1-oxy-1,2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonyl) Thioate

NlpE – New lipoprotein E

NMR – Nuclear Magnetic Resonance

PDB – Protein Data Bank

PDZ – Postsynaptic density of 95 kDa, Discs large and Zonula occludens 1

pET – plasmid for Expression by T7 RNA polymerase

PRE – Paramagnetic Relaxation Enhancement

RE – Restriction Enzyme

RseA – Anti-sigma-E factor

RseP – Regulator of sigma-E Protease

OM – Outer Membrane

Omp – Outer membrane protein

OXYL-1-NHS – 1-oxy-2,2,5,5-tetramethylpyrrolidine-3-carboxylate-N-hydroxysuccinimide ester

SEC – Size Exclusion Chromatography

Skp – Seventeen kilodalton protein

Spy – Spheroplast protein Y

ssNMR – solid-state Nuclear Magnetic Resonance

SUMO – Small Ubiquitin-Related Modifier

SurA – Survival protein A

TROSY – Transverse Relaxation Optimized Spectroscopy

Contents

Abstract	iv
Sammanfattning på Svenska	v
Publications	vii
Contribution report	viii
Abbreviations	ix
Chapter I: Introduction	2
1.1 Protein quality control	2
1.2 Inside the <i>E. coli</i> periplasmic space	4
1.2.1 Envelope stress responses in <i>E. coli</i>	6
1.3 Serine proteases	7
1.3.1 HtrA family	8
1.3.2 HtrA proteases in human	8
1.3.3 HtrA/Deg proteases in plants	9
1.3.4 Deg proteases in <i>E. coli</i>	9
Chapter II: Methodology	12
2.1 Molecular cloning	12
2.2 Protein expression	12
2.3 Protein purification	13
2.4 Outer membrane vesicles	15
2.5 Bio-layer Interferometry	15
2.6 Introduction to NMR spectroscopy	16
2.6.1 NMR relaxation	19
2.6.2 Model-Free approach	20
2.6.3 Paramagnetic relaxation enhancement	21
2.6.4 Chemical Shift perturbations	21
2.6.5 NMR diffusion experiments	21
2.6.6 Solid-state NMR	22
Chapter III: Results Discussion and Summary	24
3.1 The starting point of the projects (Paper V)	24
3.2 Probing the dynamics and interaction of DegP PDZ domains (Paper I)	24
3.3 Investigating DegP protease core (Paper II)	29
3.4 Insight to full-length DegP (Paper III)	31

3.5 Interaction between Skp and DegQ/DegP proteases (Paper IV).....	32
3.6 Concluding remarks.....	35
Chapter IV: Future perspectives	37
Acknowledgements.....	38
References	40

The Scope of the thesis

The core of the thesis is the detailed structural and dynamical characterization of the bacterial serine protease DegP by employing different NMR methods in combination with biochemical and *in vitro* assays. The thesis is split into the following four chapters:

- **Chapter I: Introduction.** The preface of the existing literature to the thesis related subjects (protein quality control, *E. coli* periplasm, HtrA serine proteases).
- **Chapter II: Methodology.** A rundown of the biochemical, biophysical and NMR methods used in the thesis.
- **Chapter III: Results discussion.** A summary of the main findings of each manuscript.
- **Chapter IV: Future perspectives.** An outline of the plans and interests for the continued work on DegP and its related protease family members.

The thesis also consists of five papers (two published articles and three manuscripts):

- **Paper I (published)** investigates the dynamics of the DegP PDZ domains over a large temperature range and reveals their crucial role in modulating the different DegP oligomeric states.
- **Paper II (manuscript)** explores the core protease domain of DegP and shows the important role of the inherent dynamics of regulatory and catalytic loops underlying DegP protease function.
- **Paper III (manuscript)** delves into the interactions between periplasmic chaperone Skp and DegQ/DegP serine proteases and elucidating proteolysis as a key step in Skp regulation.
- **Paper IV (manuscript)** is an assignment note for full-length DegP using solid-state NMR spectroscopy, providing first access to the larger oligomeric states of DegP by NMR methods.
- **Paper V (published)** is a study about α -synuclein interactions with various chaperones of eukaryotic and prokaryotic origin, leading to the identification of a key regulatory role of molecular chaperones in the physiological function of α -synuclein.

Chapter I: Introduction

This chapter introduces the concepts of the cellular protein quality control machinery with a focus on *Escherichia coli* periplasm. Furthermore, it incorporates the premise about serine proteases and in particular the HtrA family.

1.1 Protein quality control

Protein quality control is one of the main pillars of any living organism. Proteins during their lifespan are always susceptible to aggregation and misfolding, which makes them lose their functions and potentially results in toxic accumulation of protein inclusions. The misfolding/aggregation can occur due to numerous stress factors like osmotic pressure, temperature or “crowding effect”. For instance, *in vivo* protein concentrations can reach up to 300–400 mg/ml (1). Such a high protein density can have dual-effects: it can stabilize some proteins (2), but also can cause others to take on more compact non-native states that are susceptible to aggregation (3). This can lead to the destabilization of cell homeostasis or death (4). To survive and function, cells have a large arsenal of protein quality tools: molecular chaperones and proteases to control protein folding and prevent aggregation.

Molecular chaperones are proteins that can modulate protein conformation, a feature essential to many newly translated proteins *in vivo* (5). They are classified into three types: foldases, holdases, and disaggregases. All of them can be either ATP-dependent or independent (6). Hsp70 and GroEL-GroES are model examples of the ATP-dependent chaperones (Fig. 1A). Both of them use ATP as an energy source, but with slightly different mechanisms. Hsp70 binds and releases substrates using ATP (7), whereas the GroEL-GroES complex discharges substrates together with its GroES cap during ATP proteolysis, allowing for the next substrate insertion (8). On the other hand, the ATP-independent chaperones cannot use ATP as an external energy source and are, more commonly, activated under stress conditions, where they can capture and protect unfolded proteins from aggregation. These chaperones mostly have been identified in the cell compartments that lack ATP, like the periplasm in bacteria or the extracellular matrix in eukaryotes (9, 10). To give some examples, *E. coli* heat shock protein Hsp33 is instead regulated by its redox state. It is inactivated in the reduced form but is activated when oxidized (11). Another example is the periplasmic chaperone Skp that makes a stable chaperone conformation only upon binding the unfolded Omps, and delivers them for the outer membrane insertion or degradation (12).

Contrarily to chaperones, proteases can both break down the proteins or activate them by cleaving signal sequences (10). They are classified into two types: (1) exoproteases that cleave starting either from the amino- or the carboxy-terminal end, and (2) endoproteases that attack internal peptide bonds. Endoproteases have 5 sub-groups based on the catalytic site residues: serine, aspartic, cysteine/thiol, metallo, and glutamic acid/threonine (13). The serine proteases are the largest family and since quantity frequently means evolutionary success, they notably have high importance in the different organisms. To give examples, in humans, they are a part of the immune system (14) or in bacteria, they remove aggregates or unfolded proteins to prevent cell death (15). However, due to protease's destructive nature, there are either kept in the inactive form “zymogen” (16) or are expressed only during critical conditions, like apoptosis or accumulation of unfolded proteins (17, 18).

Some proteins have the dual functionality of both chaperone and protease. One of the most extensively researched examples is the protein family member Clp/Hsp100, consisting of protease core and chaperones rings that are regulated by ATP (19). For instance, both ClpA and ClpX constitute a chaperone compartment (Fig. 1B), where the substrate is first recognized and unfolded then is directly transferred to the proteolytic ClpP component for cleavage (20). The main part of this thesis is focused on the periplasmic chaperone-protease DegP, which belongs to the HtrA (protease) family. Previously described as a “death star”, DegP is a hexamer that can form complexes up to 24-mer, which makes it larger than any other major proteases or chaperones in *E. coli* (Fig. 1C) (21). Depending on temperature or other

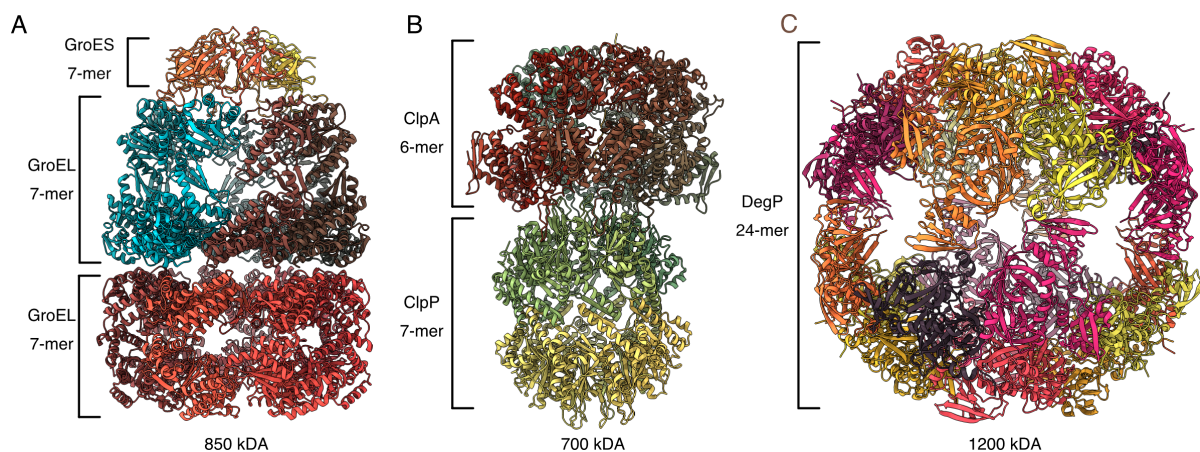


Figure 1. Architectures of *E. coli* large chaperones (A) Crystal structure of ATP-dependent GroEL-GroES, consists of heptameric cap GroES and two heptameric GroEL rings (PDB id: 1PCQ). (B) The resting state of the ATP-dependent ClpA/ClpP complex is made of two subunits: hexameric ClpA and heptameric ClpP (PDB id: 6UQE). (C) The crystal structure of the ATP-independent DegP 24-mer cage, is formed by eight DegP trimers and is one of the largest protease/chaperone complexes in *E. coli* (PDB id: 3OU0).

environmental factors, DegP can shift from chaperone to protease function to help cope with envelope aggregates and is essential for *E. coli* at elevated temperatures (22). Although this dual activity has been examined for more than 20 years and one could think that it has been completely characterized, but the key aspects, like DegP's role as a chaperone in the periplasmic space, the substrate specificity, or its temperature switch, are still vague and inconclusive.

1.2 Inside the *E. coli* periplasmic space

The cell envelope of gram-negative bacteria consists of the inner membrane and outer membrane. The space between them is called periplasm that can make up to 40% volume of the entire bacteria (23). 326 proteins are predicted to be a part of the periplasm (24) which is almost 20% of all encoded genes in *E. coli* (25). In general, the periplasm acts as an entry or exit point for many nutrients and as a protective barrier against toxic compounds (26). There is a large network of “housekeeping” proteins that oversee other proteins in the periplasmic space. Proteins destined to the periplasmic space are translated with the signal peptide and primarily translocated (across the inner membrane) by the Sec machinery (Fig. 2A). First and foremost, there are two pathways proposed for the initial nascent Sec substrate proteins. Either, SecB captures translated protein in the cytoplasm and recruits SecA or SecA interacts directly with protein during translation and afterwards may recruit SecB if it is needed (8, 27). In both cases, the protein is delivered to the inner membrane and transferred to SecYGE that translocates it further in an unfolded state (28). The signal sequence peptide is transferred first and makes a loop-like configuration, with the amino-terminus side on the cytoplasmic side and the loop in the periplasm. Once the peptide is cleaved, translocation continues (29). In parallel to the Sec pathway, there is second translocation machinery, the Tat (twin-arginine translocase) system, which, unlike Sec, translocates already folded proteins (30). The signal peptide sequences determine which translocation machinery is used (31). And whereas both sequences are generally similar, the Tat signal peptide is determined by two conservative arginines (32).

After arriving to the periplasm, the proteins are exposed to an extensive network of chaperones and proteases that do general “housekeeping” by supporting the folding, preventing aggregation, or cleaving proteins. One of the most important members to protect are outer membrane proteins (Omps), which functions vary from the non-specific diffusion pores (OmpF, OmpC) to specific transporters for iron complexes (FhuA), sugars (LamB) or are even supporting the peptidoglycan (OmpA) (33). The Omps are inserted into the outer membrane *via* the β -barrel assembly machinery (BAM) that consists of five BamABCDE proteins, from

which BamA is critical for accelerating Omp folding and together with BamD are crucial for cell survival (34). Inside the periplasm, there are two main pathways to how Omps can be delivered: SurA and DegP/Skp pathways (35, 36). As primary chaperone in the periplasm, SurA assists in the folding of outer membrane proteins like OmpA, OmpF, and LamB (37) and deletion of *surA* leads to depletion of Omp density in the membrane (38). The double knockouts *surA degP* or *surA skp* are lethal, proving that all three proteins share partially overlapped functions (39). SurA has high structural flexibility between its domains, which allows recognizing most of the Omps and coordinating binding as well as the release (40). It is assumed that periplasmic chaperones like SurA can capture Omps and give them more time to find correct structural intermediate/conformation or transport across the periplasm as shown with outer-membrane receptor FhuA (41).

The alternative pathway DegP/Skp is suggested to be a secondary route if SurA is not able to carry out its function or to remove abundant toxic Omp intermediates during the stress conditions (38, 42). DegP primarily functions as a heat-activated serine protease, even though it can capture Omps (43) and a protease deficient mutant (DegP^{S210A}) is sufficient to rescue cells at high temperatures (44). DegP shares structural and functional similarities with other bacterial DegS and DegQ proteases (Fig. 2B) (45), which can compensate in the case of DegP

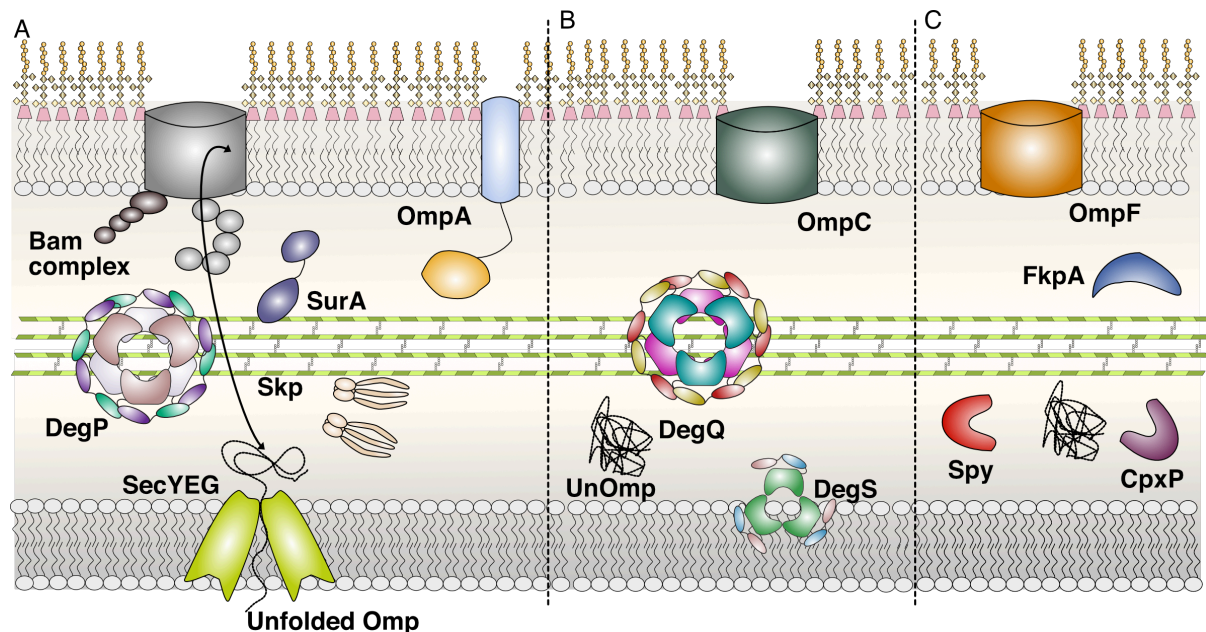


Figure 2. Overview of the periplasmic chaperones. (A) An unfolded Omp (unOmp) is transported through the SecYEG complex to the periplasmic space. SurA delivers Omp to the Bam machinery complex, which inserts it to the outer membrane. An alternative pathway for Omps, is supposed to pass through Skp and DegP, during stress conditions or when SurA is not available. (B) The homologs of DegP, hexameric DegQ and trimeric DegS proteases can recognize unOmp as well. (C) The alternative small periplasmic chaperones FkpA, Spy, CpxP can prevent unOmp aggregation in a similar manner (50, 51).

absence (46, 47). On the other hand, Skp is a holdase, that suppresses misfolding and aggregation of Omps (41). Although Skp does not fold the transmembrane parts of proteins, individual non-membrane domains of large substrates like OmpA can fold outside or inside the cavity of Skp (48). Similar to SurA, Skp also displays high flexibility, that transforms to a rigid scaffold upon binding the Omp (49).

Smaller chaperones like cradle-shaped FkpA, Spy or CpxP help as well in suppressing protein aggregation (Fig. 2C) (50, 51). FkpA is a cis/trans peptidylprolyl isomerase (52) that can substitute Skp in Omp biogenesis if present in the cell (53), whereas Spy and CpxP are chaperones that are activated by Cpx stress regulon and work together with DegP (51, 54). Other chaperones are only activated in specific stress situations, like HdeA under extremely acidic conditions (55) or for cysteine-rich proteins, the Dsb protein family helps form disulfide bonds (DsbA oxidizes disulfides, DsbC isomerizes wrongly formed bonds, and DsbD acts as a regenerative intermediate) (56). Overall, the bacterial periplasm and its management is vital for cell survival and, thus, have to be protected by different stress pathways and regulations.

1.2.1 Envelope stress responses in *E. coli*

There are five known envelope stress responses (pathways) σ^E , Cpx, Bae, Rcs and Psp that react to environmental stress (temperature, osmotic pressure), toxic compounds, or intrinsic

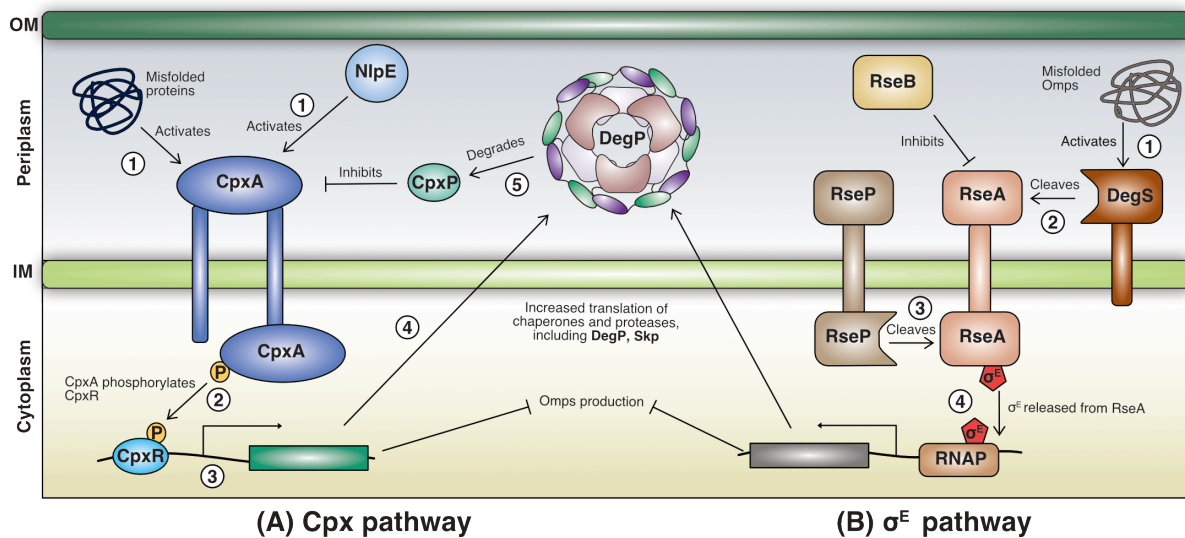


Figure 3. Two main stress envelope pathways. (A) The Cpx pathway begins *via* activation of CpxA either by misfolded proteins or the NlpE lipoprotein (1), which leads to phosphorylation of CpxR by cytoplasmic part of CpxA (2). Phosphorylated CpxR activates regulon (3) that increases translation of periplasmic chaperone and serine proteases, like DegP (4). DegP cleaves misfolded protein as well as CpxP (5), which in turn inhibits CpxA during the resting state of the cell. (B) The σ^E pathway starts with activation of DegS serine protease by misfolded Omps (1), which leads to proteolysis of RseA (2) from the periplasmic side by DegS and subsequently from cytoplasm by RseP. Cleaved RseA releases σ^E factor (4), which can then bind to RNA polymerase and kickoff regulon for increased production of periplasmic chaperones and proteases (57).

stress in *E. coli* (57). The σ^E and Cpx are critical for protein quality control. The σ^E response (Fig. 3A) is induced by the accumulation of unfolded Omps (unOmp). In the presence of unOmp, DegS cleaves the periplasmic domain of RseA (58). This allows a second cleavage to occur of RseA by the cytoplasmic RseP protease. After the complete cleavage of RseA, σ^E is released and activates transcription of its regulon, including different periplasmic chaperones and proteases (59). The second response is orchestrated by the Cpx (conjugative plasmid expression) system, a two-component system (Fig. 3B) that was first discovered *via* different phenotypes by the cause of mutations in the *cpx* locus (60). Cpx sensing ranges from misfolded envelope protein to intermembrane proteins. Identified as a two-component signal transduction system, the *cpx* locus consists of *cpxA* and *cpxR* genes (61). CpxA acts as kinase and phosphatase for the response/transcriptional regulator CpxR, and is inhibited by the chaperone CpxP. Phosphorylated CpxR activates the transcription of periplasmic chaperones and proteases (62). One of the activated proteases is DegP, which in turn degrades CpxP and thus further stimulates CpxA activation (50). *degP* expression is upregulated by both, σ^E and CpxR, indicating that DegP is the primary protease under envelope stress (63). The remaining three systems are activated by toxic compounds (Bae), LPS defects (Rcs) or help maintain the proton motive force and take care IM protein misfolding (Psp) (64).

1.3 Serine proteases

Serine proteases are one of the most abundant enzymes in the human genome. They make up about one third of all known enzymes. On that account, they are involved in many different cellular processes such as inflammation, cell differentiation, apoptosis, or general protein quality control (14). They are classified into 13 clans and 40 families based on the catalytic residues and amino-acid specificity (65). The largest clan is PA, which consists of two sub-clans (SA, CA). Consequently, they are based on the utilization of either serine (SA) or cysteine (CA) as nucleophilic residue. The PA clan is mostly found in eukaryotes and the remaining clans are more represented in other organisms (66). The catalytic triad of SA clan usually consists of histidine, aspartic acid, and serine. Serine acts as a nucleophile, which attacks peptide bonds through acyl-enzyme intermediates. Histidine acts as a base during the cleavage, where it gets protonated and afterwards stabilized by aspartic acid through a hydrogen bond (67). The specificity of serine proteases substrates is defined by the surface loops that are called (using chymotrypsin nomenclature) LA, LB, LC, LD, LE and L1, L2, L3 (68). Loops L1–3 usually contain catalytic residues and help form an active center, whereas loops A–E perform allosteric or inhibitory functions (69).

1.3.1 HtrA family

Among all the serine protease families, the high-temperature requirement A (HtrA) family is known for a highly conserved sequence between different organisms that can range from bacteria to human cells (Fig. 4A). HtrA belongs to SA protease clan in which proteins commonly have a two-domain structure with each domain forming a six-stranded β barrel (70). HtrA proteins can be easily identified from other serine proteases due to the presence of one or two PDZ (postsynaptic density of 95 kDa, Discs large and zonula occludens) domains (71). An exception is Deg proteases from plants that can have none or up to 4 PDZ domains, however, they are the least studied members of the HtrA family (72). The best known HtrA proteins are the bacterial proteases DegQ, DegS, and DegP, which coordinate protein quality control in the periplasmic space and they serve as structural and functional models for the whole HtrA family (73).

1.3.2 HtrA proteases in human

Four HtrA proteases called HtrA1, HtrA2, HtrA3 and HtrA4 have been identified in humans. They have 50–60 % sequence similarity, but their structural conformation is quite different from each other (74). HtrA 1, 3, and 4 have very unique amino-terminal domains consisting of IGFBP and Kazal-like modules, but their importance remains still largely unknown (75). They are involved in arthritis, cancer, neurodegenerative disorders, and potentially other pathological diseases (76). HtrA1 forms trimers similar to other HtrAs, but in contrast to them, the PDZ domain is not required for protease activation and this implies that there are different specific regulatory mechanisms (77). HtrA2 is found in mitochondria and when it is released from it, it induces cell death by removing the apoptosis inhibitors of caspases -3, -7 and -9. (78, 79). Structurally, HtrA2 adopts a typical trimeric form, where the PDZ domain is essential for protein-protein interactions (80). HtrA3, due to alternative splicing, is expressed in two isoforms, which are called long (HtrA3L) and short (HtrA3S), with the latter lacking its PDZ domain (81). The HtrA3 structure is very similar to the HtrA1 and its PDZ domain is also dispensable for proteolytic activity, but necessary to form trimeric structures unlike HtrA1 or HtrA2 (82). Finally, there is not much known about HtrA4, but lately, it has been shown it can degrade apoptosis inhibitor XIAP like HtrA2 and has potential interaction partners like caspases 9 and 7, cytoskeleton proteins, TCP1 α chaperonin, and calcium-binding S100A6 protein (83).

1.3.3 HtrA/Deg proteases in plants

Plants possess a much bigger variety and diversity of Deg proteases compared to other kingdoms. In *Arabidopsis thaliana* alone there have been 16 *deg* genes identified and Deg proteases are found in almost all compartments of the cell: lumen/stroma of chloroplasts, mitochondria, peroxisomes, and the nucleolus (84). The solved crystal structure of Deg1 and Deg2, both proteins are involved in the repair/degradation of photosystem II during light stress (85, 86), revealed that they form hexameric structures similar to (bacterial) DegP/DegQ, and that their activity is heavily dependent on the surrounding pH. On the other hand, the Deg9 protease forms octamer that consists of two tetrameric rings, where one has active and the second non-active conformations (87). On top of that, it is worth mentioning, that Deg proteases have been identified in cyanobacteria too, however, their localization and functions related to photosynthesis are still unknown (88).

1.3.4 Deg proteases in *E. coli*

To current date, three HtrA proteases have been identified in *E. coli*: DegP, DegQ, and DegS. All three proteases target hydrophobic sequence elements, particularly next to valine or isoleucine residues (89). DegP and DegQ are the closest homologues, both contain two PDZ domains and are hexameric in cellular ground states. After binding the substrate, they can form either 12- or 24-meric structures and are debated to possess both protease and chaperone

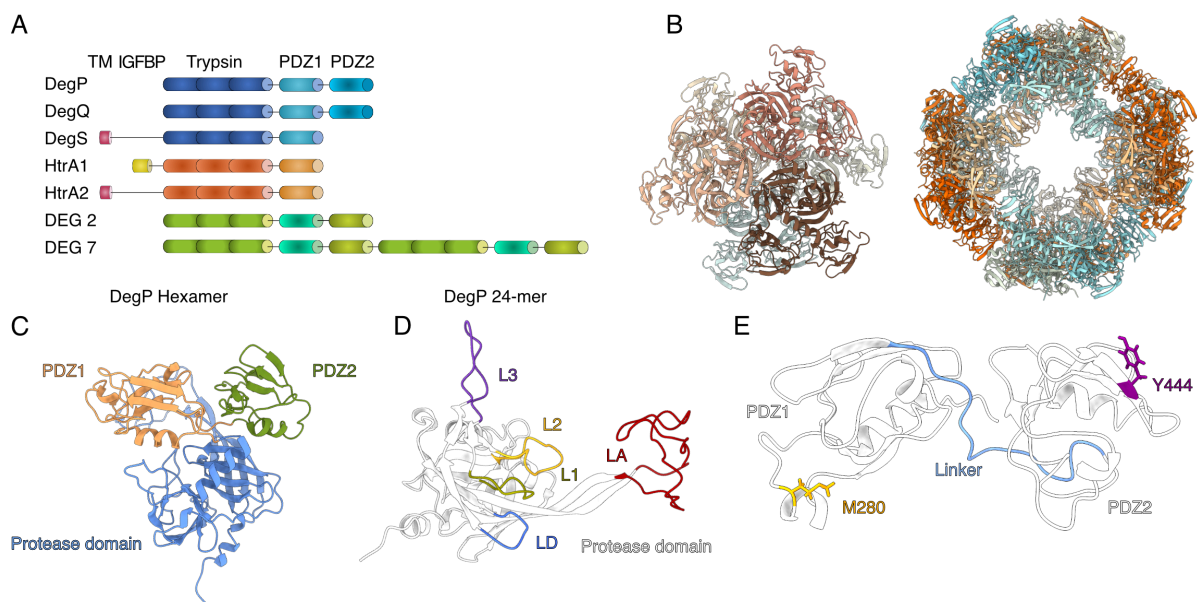


Figure 4. The architecture of the DegP serine protease. (A) Crystal structure of DegP hexamer (PDB id: 1k9j) and DegP 24-mer (3OU0). (B) HtrA protein family and their domain classification. TM stands for transmembrane sequence, IGFBP for insulin-like growth factor binding protein (70). (C) Domain architecture of DegP monomer. (D) DegP protease domain loop arrangement. (E) DegP PDZ1-PDZ2 domains structure and cage-linking residues M280/Y444.

functionality (Fig. 4B) (90). It has been observed that DegP is allosterically activated by hydrophobic carboxy-terminal residues (91) and primarily protects cells during heat-stress (92), whereas DegQ function is more related to pH changes in the cell envelope (93). Furthermore, there are also reports of DegP extracellular activity, where it inhibits biofilm formation of pathogenic bacteria (94) or that it can cleave E-cadherin, which weakens the entry barrier to the epithelial cells (95). This has been associated with several pathogenic bacteria along with DegP being potentially a part of diseases such as Lyme, diarrhoea, sepsis, salmonellosis and a few more (96).

The first crystal structure of DegP revealed that it is a hexamer of two trimeric units. Protease domains are in the core, surrounded by PDZ domains, that work as gatekeepers for the substrate (Fig. 4C) (97). In later studies, it was observed that the inactive hexamer and the 24-mer of DegP assumed different domain arrangements. The DegP hexamer is stabilized *via* interactions of PDZ1-PDZ1 domains, whereas in the substrate-bound form it changes to PDZ1-PDZ2 interaction and reorientation of L3 loop to PDZ1 domain (98). It was shown that model substrates bind to a PDZ1 pocket and activate DegP (99). This led to the proposal that PDZ1 interaction is an essential step for initiating proteolysis, however, later structures proved the substrates without PDZ binding degron, can still activate DegP, implying the existence of different activation pathways (100). Nevertheless, the next step in the activation cascade happens in the DegP loop system, which consists of regulatory LA, LD and active centre controlling L1 L2, L3 loops (Fig. 4D). The LA loop is an inhibitory loop, which unwinds from L1 and L2 upon increased temperature for enhanced proteolytic activity (101). The L1 loop contains the active site serine and L2 is critical for substrate binding (97, 102). L3 acts as a sensor loop for PDZ1 and due to the PDZ1-triggered allosteric activation, it can reorientate to the protease domain for both activating and inhibitory effects (90). Lastly, the LD loop is transmitting an allosteric signal *via* interaction with other loops (103). The proposed DegP loop activation cascade is L3-LD-L1-L2, where L3 loop senses the activation signal and broadcasts it to Arg178, which forms interactions with LD Leu174 and Glu175 residues. The LD conformation changes transmit to L1 that in turn sets up catalytic triad (98).

The substrate-bound DegP oligomers have been extensively researched due to a great interest in their enormous size that dwarfs even the GroEL chaperone (21). The inner cavity of DegP 24-mer is 110 Å long, that can easily hold even the largest Omps (104). Surprisingly, cage complexes are not required for proteolysis or cell survival but may contribute to general cell fitness by suppressing unnecessary proteolytic activity (105). The main cage-oligomers are 12-

mer and 24-mer size, but 9-mer oligomers have been observed as well by mass spectroscopy, however, only 9 substrates could occupy them, where 12-mer can bind up to 24 peptides (106). The DegP cages are stabilized *via* packing of PDZ1 domain in one trimer with PDZ2' domain from the adjacent trimer. The key residues in the trimer interface are L276, M280, F289 from PDZ1 and V431, Y444, L446 from PDZ2' domains, where Y444A mutation abolishes the formation of the DegP cages (Fig. 4E) (105).

The last member of bacterial HtrA proteases, DegS possesses only one PDZ domain and it is anchored to the inner membrane *via* its C-terminal end/terminus. Contrary to DegP or DegQ, DegS is exclusively found as a trimer. DegS is a sensor of envelope stress by recognizing and binding unassembled Omps, upon which DegS is activated and subsequently cleaves RseA that leads to the induction of the σ^E regulon (58, 107). Furthermore, DegS recognizes similar peptides as DegP, suggesting that they might share substrates despite structural differences (108).

Taken together, HtrA proteases play an active role in cell defences with parallel and distinct functions at the same time. Considering, that DegP is the most scrutinized of all of them, it was, naturally, best suited for a starting point of HtrA proteins research by NMR spectroscopy.

Chapter II: Methodology

This chapter covers the overview of methods used in the thesis with examples and comments for motivations of choosing them. For the experimental details, the reader is referred to the corresponding papers.

2.1 Molecular cloning

The DNA constructs for the thesis were prepared using restriction enzyme (RE) cloning and site-directed mutagenesis or bought directly from GenScript. For the (RE) cloning, the genes of interest were PCR amplified with primers from the *E. coli* genomic DNA or the plasmid. The purified empty vector and the purified PCR product (insert) were then digested with respective restriction enzymes according to the manufacturer's instructions. The cleaved and purified vector and insert were ligated with T4 DNA ligase and transformed to the competent cells (109). The site-directed mutagenesis was done by designing partially complementary primers with the desired mutation in the middle and performing PCR. The product was subjected to DpnI restriction (target site 5'-Gm6ATC), to selectively remove the original template, which is methylated. After digestion and inactivation of the enzyme, the product was transformed into XL gold competent cells, which have very high efficiency for accepting DNA. Afterwards, several colonies were selected for plasmid extraction and sequencing (110, 111).

The full-length genes used in this thesis (*degP*, *degQ*, *skp*, *surA*, *hsc70*) were cloned with an amino-terminal hexa-histidine (6xHis)-tag, but for individual domains constructs (PDZ1, PDZ2, PDZ1-PDZ2), the respective gene was fused with 6xHis-SUMO (small ubiquitin-like modifier) tag. The advantage of the SUMO-tag is that it enhances the solubility of the expressed recombinant protein and is easily cleaved off by SUMO specific proteases, without leaving any additional residues at the amino-terminus of the fusion-protein / recombinant protein (112).

2.2 Protein expression

All proteins were expressed in *E. coli* BL21(DE3) harbouring an inducible T7 polymerase under the *lac* promoter in the genome (113). In the absence of lactose, the *lac* promoter is inhibited by the *lac* repressor, which binds to the operator region. When lactose enters the cells, it is converted to allolactose by β -galactosidase. Allolactose binds to *lac* repressor and reduces affinity for *lac* operon, thus allowing T7 RNA polymerase binding to operon (114). Non-labelled proteins were expressed in LB-medium induced by the non-hydrolyzable lactose analogue IPTG (115).

For isotope labelling, proteins were grown either in water-based or deuterated M9 (minimal) medium supplemented with $^{15}\text{NH}_4\text{Cl}$ and (^{13}C) -glucose or $(^{13}\text{C}^2\text{H})$ -glucose, respectively (Fig. 5A) (116). Deuterated proteins are used to suppress spin diffusion due to lowering the density of protons in the protein. Likewise, deuteration decreases relaxation rates of ^{13}C since carbon is directly bonded to deuterium and deuterium spin has a lower gyromagnetic ratio. Overall, deuteration helps with improving line bandwidth and signal-to-noise ratio in NMR measurements for large proteins (117).

For methyl-group labelling of isoleucine, leucine and valine (ILV), cells are cultured in completely deuterated M9 (minimal) medium and selectively protonated amino acid precursors, $[\text{H}]$, 3,3- $[\text{H}_3]$ -ketobutyrate and 3- $[\text{H}]$, $[\text{H}_3]$ -ketoisovalerate, are added to the culture one hour prior to induction of protein expression (118). As these compounds are precursors for ILV synthesis in *E. coli*, they are incorporated directly without scrambling to other amino acids as these amino acids are end points of the biochemical pathways. To label the methyl-groups of alanine and methionine, selectively protonated amino acids, 2- $[\text{H}]$, 3- $[\text{H}_3]$ -L-alanine (119) or $[\text{H}]$, $[\text{H}_3]$ -L-Methionine (120) are added before induction. The advantage of the MALVI-labeling is that it allows the detection of methyl groups, which give strong and well-separated signals in $[\text{H}_3, \text{H}]$ -HMQC spectra, even observable for proteins and protein complexes up to 1 MDa (121, 122).

2.3 Protein purification

Most of the proteins were purified using established techniques such as immobilized metal affinity chromatography (IMAC), ion-exchange chromatography, and size-exclusion. IMAC was introduced in 1975 by Porath and his colleagues, by making gel beads with immobilized metal chelating ligands (123). Proteins though histidines or cysteines can bind to them *via* the coordination bonds of divalent metal ions and by using varying concentrations of competing organic compound – imidazole, the proteins can be eluted in different fractions. Hence, by introducing a tag of six or more histidines to the amino-terminus or carboxy-terminus of target proteins, it was made possible to enhance protein binding to the beads and purify different proteins with higher purity using IMAC (124). In addition to histidine tags, an additional tag can be attached, like MPB (maltose-binding protein), GST (Glutathione-S-Transferase), SUMO, etc.. Considering that two steps of the IMAC purification are required when using these tags: one for eluting full protein with a tag and the second for removing a cleaved tag from protein, it greatly improves purity. Additionally, tags help in stabilizing the target protein in the initial steps of purification and can help in increasing solubility during expression (Fig.

5B) (125). The next technique widely used for purification is ion-exchange chromatography. As the name suggests, it is based on the surface charge of the protein and the matrix. Proteins with a positive charge bind to a negative matrix and *vice versa*. The protein charge depends on the pH of the buffer and the isoelectric point (pI) of the recombinant protein, if pH is higher than pI, the protein charge will be negative if it is lower – positive (126). The last method – gel filtration is based on the separation of large and small molecules/proteins due to their size by reason of larger proteins passing through the beads whereas smaller ones get stuck to the small pores in the beads (126, 127).

Besides the standard purification methods, the refolding procedure was used along with certain proteins (DegP, Hsc70, Skp, SurA). The chaperones and, especially, native *E. coli* proteins frequently are co-purified with substrates or small peptides (49, 104). From an NMR perspective, refolding is also important for deuterated proteins as they will have deuterated

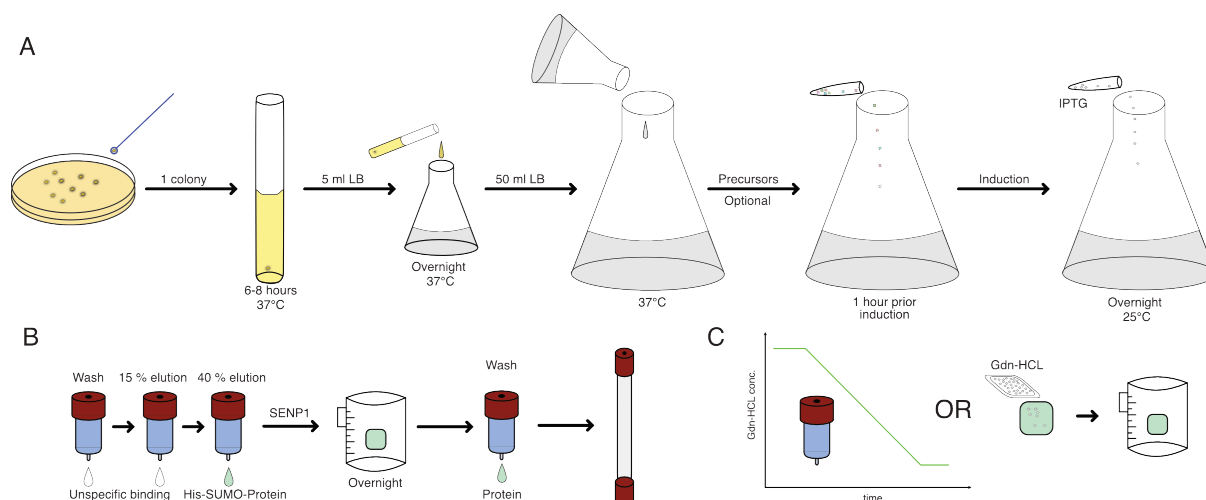


Figure 5. A rundown of protein production. (A) Isotope labelled protein expression scheme. A single colony from freshly transformed cells is inoculated in 5 ml of LB medium and incubated at 37°C at 70 rpm for 6-8 hours. Subsequently, 1 ml of preculture is diluted in 50 mL of freshly prepared M9 medium and incubated at 37°C and 75 rpm overnight (o/n). The o/n culture is then diluted in 950 ml M9 medium and incubated at 37°C and 75 rpm until the main culture reaches an OD₆₀₀ (optical density at 600 nm) of 0.6-0.8. Recombinant protein expression is then induced with 0.4 mM IPTG and the culture is incubated at 25°C and 75 rpm overnight (approx 16 h). (B) Simplified His-SUMO tag purification procedure. The cell lysate is loaded to Ni²⁺ column and subsequently washed with wash buffer, followed by 15% and 40 % elution buffer. The target His-SUMO tagged protein is eluted with 40% elution buffer. Afterwards, the protein is dialysed against PBS and simultaneously incubated with SENP1 protease for cleavage of the His-SUMO tag overnight. Finally, protein is loaded to Ni²⁺ column again and flowthrough is collected with target protein. As a final purification step, protein is concentrated and applied to the gel-filtration column. (C) Refolding strategies. Refolding protein on-column by reducing Gdn-HCl concentration via gradient or adding Gnd-HCL directly to protein solution and dialysing it out overnight.

amide moieties in the protein core, which will not give any NMR signals if they are not back exchanged to hydrogens (128). For the proteins used in this thesis, DegP and Hsc70 were both refolded on the column, whereas Skp and SurA were refolded using dialysis (Fig. 5C) (129–131).

2.4 Outer membrane vesicles

Outer membrane vesicles (OMVs) are emitted from gram-negative bacteria periplasm during cell growth. They are approximately 10 to 300 nm diameter spheroid particles and as they leave bacteria, they entrap some of the periplasm together with proteins inside (132). While the origins of OMV regulation are still under investigation, they are already utilized for native proteins studies (133). OMVs can be enriched by periplasmic proteins by overexpressing them and collecting discharged OMVs from the medium (134). The OMVs from the wild-type strain B21(DE3) contains a lot of native Omps, which makes it difficult to apply any assays that require high purity protein, however using B21(DE3)omp8 that lacks major Omps, improves entrapped protein purity up to 90% (135). By employing this method, the OMVs have been used as a platform for studying by NMR the periplasmic expressed CpxP and MalE proteins in their native environment. The results showed well-resolved spectra and are expected to advance envelope protein studies (136).

2.5 Bio-layer Interferometry

Label-free technologies are often sought for protein characterization and interactions studies. One such method is bio-layer Interferometry, which was used to confirm interactions and as a complementary to NMR experiments. BLI is based on reflectometric interference spectroscopy: the light beam is shot to the interference layer and the wavelength of the reflected beam is measured. If the layer is free of any analyte, the wavelength should be the same as the initial beam. However, if the analyte binds to the layer, its thickness will change and with it the reflected beams wavelength. The wavelength shift will be equal to binding strength (137). In practical use, the layer is a biosensor tip, where different proteins can be immobilized based on the choice of the sensors (His-tag, Streptavidin, Protein A). When sensors with immobilized proteins are dipped to analyte solution, we observed the same process happens, the binding of analyte will change tip thickness and based on binding strength, we can determine kinetic values: association (k_a) and dissociation (k_d) rates, equilibrium dissociation constant (K_D) (138). In this thesis, I used BLI to confirm interactions between DegP individual PDZ1 and PDZ2 domains.

2.6 Introduction to NMR spectroscopy

Nuclear magnetic resonance (NMR) is a classical spectroscopic technique for observing different molecules and proteins. The beginning of NMR is considered 1946 year when F. Block and E. M. Purcell demonstrated nuclear magnetic resonances in condensed matter, for which they were awarded the Nobel prize in Physics in 1952. The main aspect of NMR is that it is possible to observe changes in the Larmor frequency of a nucleus due to the chemical bonding of the atom. Larmor frequency describes precessional motion of active nucleus rotation in the magnetic field. This allows identifying specific atoms and their different properties as well as interactions. Of course, NMR can be only applied to active nuclei that have a spin, a magnetic moment. Nuclei with even numbers (^{12}C , ^{16}O , ^{32}S) of protons and neutrons do not possess spins, whereas the ones with odd numbers (^1H , ^{13}C , ^{15}N , ^{19}F) do (139). The NMR active nuclei have two states, that are equally distributed, but in the magnetic field, the ratio of states changes to the Boltzmann distribution. The difference between the two states is equal to a certain energy level, which is proportional to the gyromagnetic ratio of the element and the magnetic field strength. Since all spins are surrounded by different magnetic environments, they will experience slightly different local magnetic fields and will have different energy levels (Fig. 6A). Due to that, the resonance frequencies in spins will be different compared to the applied magnetic field. This difference is referred to as the chemical shift.

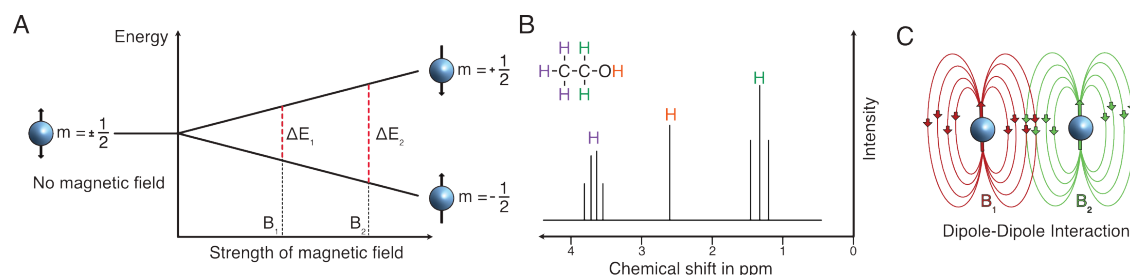


Figure 6. NMR theory. (A) The spin splitting in different magnetic fields (B_i) and its energy levels (E_i) is based on the strength of the magnetic field (B_1 , B_2) (B). 1D ^1H NMR spectrum of ethanol. The colour of protons indicate to which signals they belong. (C) Dipolar interactions between two spins.

Physically, inside the NMR spectrometer, the sample is put within the coil, which induces a homogenous magnetic field. Different radio frequencies pulses are applied to the sample and the resulting oscillating electric currents are detected that are called free inductions decay (FID). FID *via* Fourier transformation is converted to the signals (140).

The basis of protein NMR is to assign the observed signals (resonances) to specific residues. For small molecules, it can be done on the 1D ^1H spectrum, by knowing proton splitting patterns and intensities, which correlate proportionally to the number of protons (Fig. 6B) (141). However, this is hardly possible for proteins due to them having a high number of protons and structural variations. For them, it is mostly done by correlating amide resonances to different carbons atoms of selected (C^α , C^β , C) and preceding ($\text{C}_{\alpha-1}^\alpha$, $\text{C}_{\beta-1}^\beta$, $\text{C}_{\alpha-1}$) residues (Fig. 7A). This is possible due to scalar coupling. The scalar coupling or J coupling occurs between two nuclei that are connected *via* a chemical bond. The J coupling provides information about

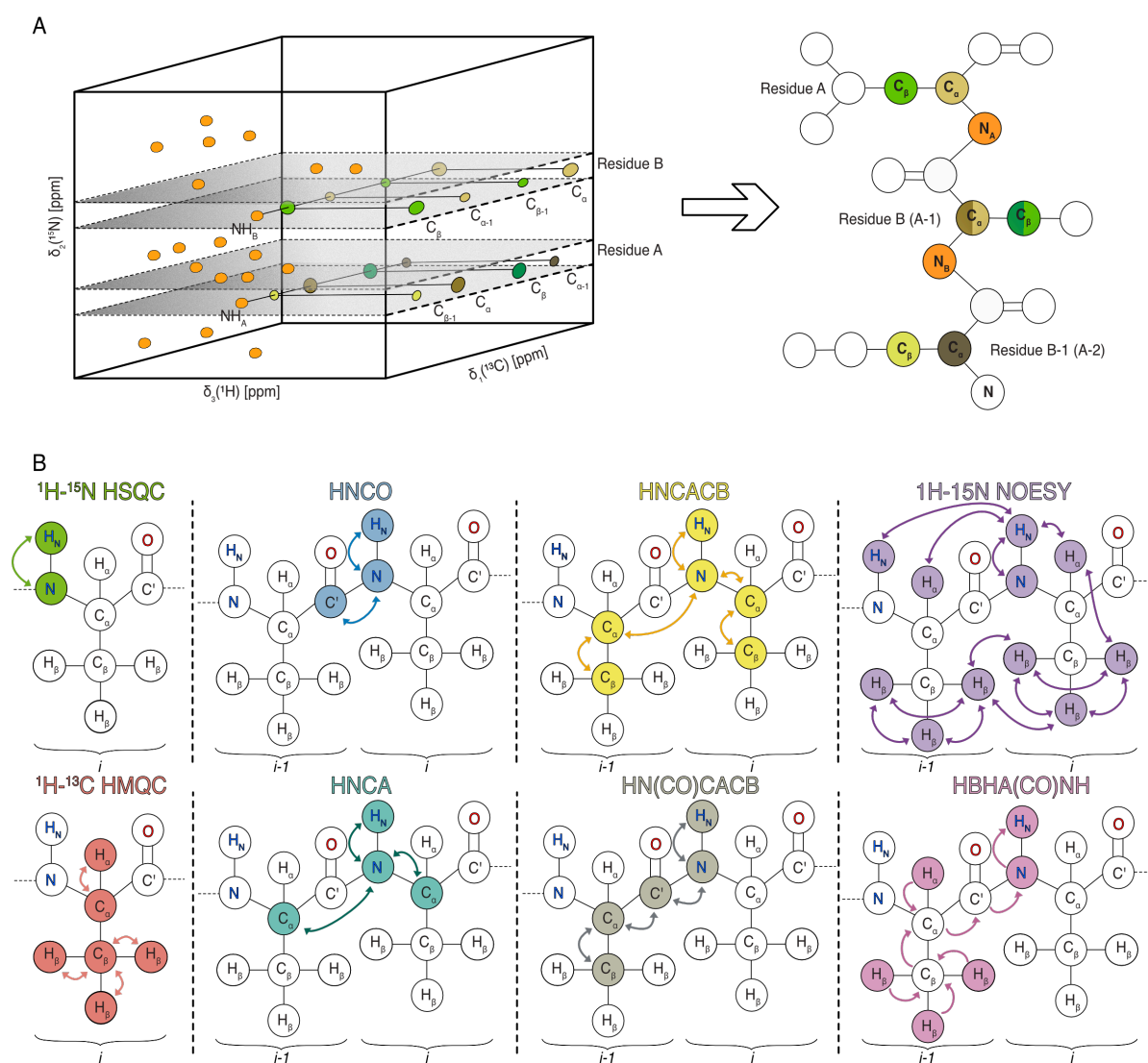


Figure 7. Standard NMR experiments. The three-dimensional measurements of double-labelled (^{15}N , ^{13}C) protein (A). Each resonance on ^{15}N , ^1H plane has signals in ^{13}C dimension and correspond to certain atom in amino acid residues. The magnetization transfer in different NMR experiments (B). Colored atoms indicate the detected resonances.

bond connectivity and can be exploited for magnetization transfers in multidimensional experiments (142). The typical experiments for backbone assignments are ^{15}N - ^1H HSQC, 3D HNCO, 3D HNCA, 3D HNCACB, etc. (the first letter of 3D experiments indicate magnetization transfer start and end) (Fig. 7B). Afterwards, it is possible to continue with the rest carbon atoms to assign side-chains. For the structure determination, it is necessary to record Nuclear Overhauser Effect Spectroscopy (NOESY) experiments (143). The NOE is caused by the direct, through space-detected magnetic interactions (known as dipole-dipole couplings (DD) (Fig. 6C) due to cross-relaxation of the spins. Hence, NOE gives information about the spatial proximity of different nuclei (144, 145). However, difficulty in identifying NOE connections increases with the protein size and for getting quick information about protein secondary structure, there is a simpler method. It is known that C^α , C^β atoms experience a downfield chemical shift in α -helices and upfield shift in β -strand (146, 147). Therefore by comparing carbon chemical shifts with random coil carbon shifts, which can be predicted using computational methods with high accuracy (148), we can identify the secondary structure elements.

Depending on the protein, NMR measurements and analysis can take a lot of time, therefore during the last 20 years, non-uniform sampling (NUS), automated assignments and structure determination have become a major part of routine tasks, that can save months or years of manual labour (149). For the experimental side, it might take days or months to record multi-dimensional spectra due to required points in indirect dimensions. This issue is mostly addressed with the help of NUS. During NUS measurement, only partially (10–25%) data is recorded and the rest is reconstructed with mathematical algorithms like compressed sensing (CS) (150, 151). CS is based on the idea, that most of the data we record is “thrown away” anyway and it is just enough to record important parts (152).

The automated analysis is frequently done by CYANA software. The CYANA has an FLYA algorithm, that capable of assigning backbone sequences with high accuracy if a sufficient amount of data is provided (149). The algorithms for the solid-state (ssFLYA) (153) and side-chain methyls groups (MethylFLYA) (154) assignments are included as well. The structure is calculated by providing distance constraints, which were obtained from NOESY spectrum. Based on NOESY cross peak volume, CYANA assigns distance range and uses torsion angle dynamics to calculate the optimized structure. (155).

2.6.1 NMR relaxation

We often image proteins as static structures and that the folded state is the final structure, however inherently they are governed by their dynamic motions (156). Protein fluctuations and energy landscapes are complex and unavoidable, thus requiring sensitive methods to observe them (157). Favourably, NMR allows us to observe a wide range of motion timescales: from picosecond to millisecond and to determine protein different conformations along with excited states (158). Traditionally, protein backbone dynamics are measured by a set of relaxation experiments, which includes longitudinal relaxation (T_1 or R_1), transverse relaxation (T_2 or R_2), and steady-state heteronuclear NOE (hetNOE). T_1 , also known as spin-lattice relaxation, describes the recovery rate to z component to the equilibrium (Fig. 8A). T_2 , spin-spin relaxation, occurs due to decoherence of spin magnetization at the x-y plane and details its return to equilibrium (Fig. 8B) (159). Both T_1 and T_2 are affected by chemical shift anisotropy (CSA). The electrons surrounding the nuclei are not distributed equally as a sphere, which means chemical shifts are anisotropic and influenced by the orientation of nuclei together with their positioning to the magnetic field (160). In solutions, it is hardly noticeable, but in solids or for large proteins the CSA becomes a large factor for T_1 and T_2 relaxation. The hetNOE

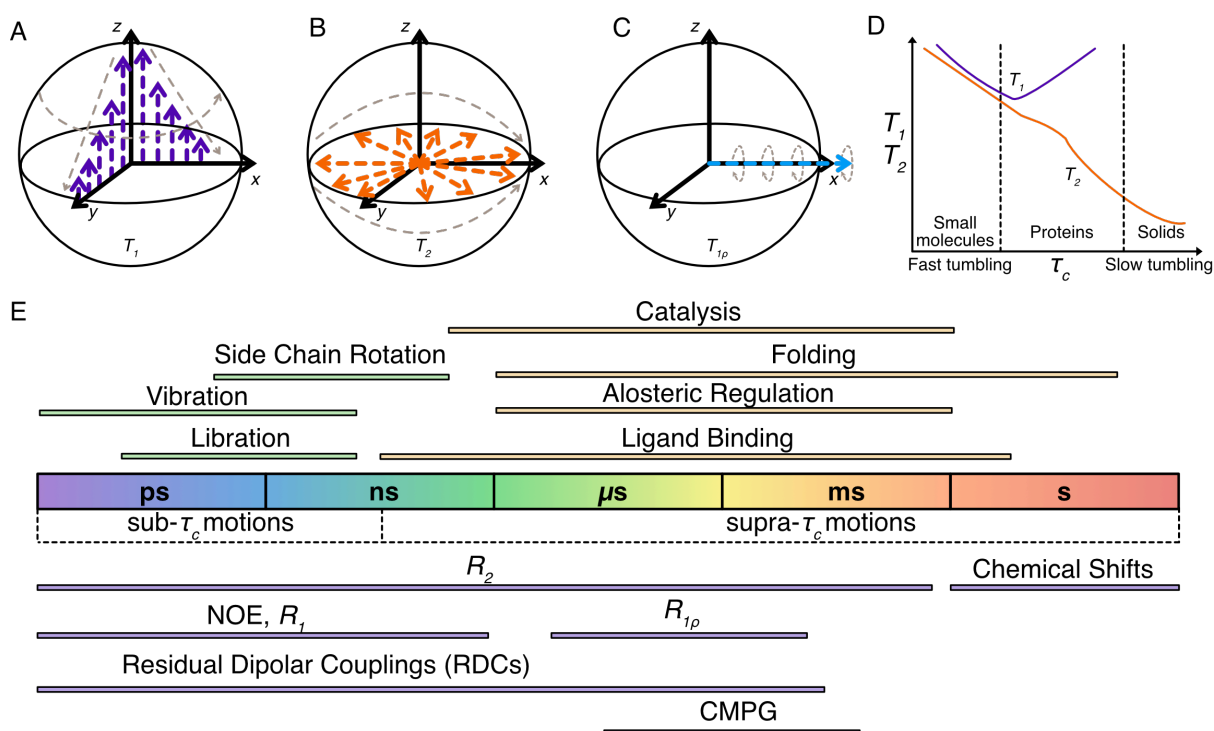


Figure 8. Relaxation overview. (A) Magnetization vector relaxation across the z plane in T_1 , (B) along the x-y plane in T_2 and (C) in rotating frame in $T_{1\rho}$ experiments. (D) T_1 and T_2 correlation to τ_c . (E) Relaxation and chemical reaction timescales. (E) the panel adapted from (159, 163).

describes relaxation parameters that are faster than the protein tumbling and is determined through a ratio of saturated longitudinal polarization of ^{15}N spin under ^1H irradiation and longitudinal polarization of ^{15}N spin at the thermal equilibrium. Although the hetNOE is crucial for backbone dynamics, its sensitivity is considerably lower than T_1 or T_2 , due to the ^{15}N spin magnetization start rather than ^1H (161).

T_1 and T_2 are coupled to protein size and their tumbling. The larger the protein, the slower the T_2 relaxation rates will be, which will result in line broadening (Fig. 8D). For many early years of protein NMR, this was the largest obstacle, but in the late twenties, Transverse Relaxation-Optimized Spectroscopy (TROSY) was introduced. TROSY is based that in magnetic fields, the T_2 relaxation is dominated by the CSA and DD. CSA and DD interfere with spins and gives different relaxation rates to spin components. TROSY focuses only on one of the components consequently removing effects of different transverse relaxation effects (162).

For studying slow processes (μs and ms) $T_{1\rho}$ or Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersions are often used. In CPMG, a train of 180° pulses is applied, which refocuses magnetization. After each pulse, signals are phased back and produces an echo signal, whose magnitude is correlated to relaxation rates (163). Slow motions also affect isotropic chemical shifts, which contribute to transverse relaxation. Therefore, with CPMG and $T_{1\rho}$ one can quantify chemical exchange and it is especially, useful for large proteins due to the high sensitivity of methyl groups (although, it can be applied for any NMR active nuclei) (163,164).

Taken together, relaxation is NMR spectroscopy true power, that allows observation of dynamic properties of proteins. With a massive NMR relaxation experiments library, it is possible to monitor real-time biochemical reactions, allosteric pathways, folding mechanisms and many more biological events (Fig. 8E) (165).

2.6.2 Model-Free approach

Model-free approach was the first time described by Lipari and Szabo (166) and is used to extract internal motions of proteins from the globular tumbling. Several parameters are obtained from MF analysis: a generalized order parameter (S^2), which is a generalized parameter reflecting the amplitude of motion, the effective internal correlation time (τ_e) that corresponds to motions faster (picosecond to nanosecond) than overall molecular tumbling, and chemical exchange rate (R_{ex}), which describes contributions to the transverse relaxation rate due to micro-mili second motions (166, 167). With these model-free parameters, five model are constructed: $\{S_2\}$, $\{S_2, \tau_e\}$, $\{S_2, R_{\text{ex}}\}$, $\{S_2, \tau_e, R_{\text{ex}}\}$ and $\{S_2^2, S_2^2, \tau_s, \tau_s\}$. The last model

includes the internal motions in faster (S^2_f , τ_f) or slower (S^2_s , τ_s) timescales. The best model is chosen *via* repeated fitting and assigned to each residue (168). To obtain these values, a set of relaxation data (R_1 , R_2 , hetNOE) has to be recorded at two different magnetic fields due to individual magnetic field intrinsic properties. Over the years several software tools were developed to help and streamline the analysis of MF data: Tensor2 (169), FAST-Modelfree (170) and relaxGUI (171).

2.6.3 Paramagnetic relaxation enhancement

Over the last two decades, PRE became one of the commonly used NMR techniques to investigate dynamic properties in large proteins (172). By attaching a paramagnetic spin-label, nearby nuclei signals are broadened due to enhanced transverse relaxation rates (173). The PRE method is very useful for determining structural distances in protein (49) or measuring the chaperone-substrate complexes' lifetime (49). There are two types of probes used for PRE: the nitroxide stable radical and metal chelators that bind paramagnetic metals. Both types have to be chemically attached to protein or nucleic acid (174). In this thesis two types of spins labels were used: nitroxide spin label MTSL which attaches to a cysteine *via* a disulfide bridge and (175) and OXYL-1-NHS that labels ϵ -amino groups of lysines (172).

2.6.4 Chemical Shift perturbations

Chemical shift perturbations (CSP) or complexation induced changes in chemical shift (CIS) are experiments to identify protein-protein and protein-ligand interactions. Typically, an unlabelled protein is titrated to ^{15}N labelled protein and for each titration step a [^{15}N , ^1H]-HSQC spectrum is recorded. Resonances that experience changes in their magnetic environment will shift in the spectrum, due to a new interaction or structural changes in the protein. In the event of the fast exchange, a signal peak move until the interaction is saturated, but during slow exchanges, the initial signal will get weaker with a new one emerging in a different position (176). CSPs are very sensitive to changes in the chemical environment, which can occur even due to buffer or salt differences (177). However, in cases, there are strong interactions or with a large protein, the chemical shifts will be broadened beyond detection due to enhancing transverse relaxation rates, but, fortunately, they can still be used to assess the binding site (178).

2.6.5 NMR diffusion experiments

NMR diffusion was used initially as a tool for measuring chemical compounds' diffusion coefficients (179). Based on the Stokes-Einstein equation, the diffusion coefficient (D) is

inversely proportional to the molecule's hydrodynamic radius and viscosity of the solution, hence measuring D , allows estimation of molecules' (e.x peptides, proteins) size (180). Further improvements, allowed the development of Diffusion Ordered Spectroscopy (DOSY) that displayed chemical shifts and diffusion coefficients on the same projection (181). However, there were several issues with diffusion spectroscopy, since diffusion gradient pulses tend to induce eddy currents in surrounding metal objects. In turn, eddy currents produce magnetic fields that distort measurements (182). To minimize these effects, we used a Bipolar Pulse Pair Longitudinal Eddy current Delay (BPPLIED) pulse sequence, which improves on DOSY, by introducing closely spaced, short gradient pulses and bipolar gradients, with different polarities, but the same magnitude. The eddy currents induced by opposite polarities, cancel each other out (179, 183).

2.6.6 Solid-state NMR

Solid-state nuclear magnetic resonance (ssNMR) is a technique to study solid or immobilized (sedimented) samples. In a solid state, the spins are orientated randomly, which gives each spin a different spectral frequency, resulting in overlapping and broad spectra. As this was not very informative, new techniques were desired to get high-resolution NMR spectra. One of such techniques was magic angle spinning (MAS), which became a routine method for the majority of ssNMR experiments (184). During MAS, the sample is in uniaxial rotation at the angle of 54.74° , otherwise known also as the “magic angle”. During this rotation angle, the chemical shift anisotropy is close to zero, which is the main problem of line broadening in solid-state powder samples (185). The reason is that in solution NMR, CSA or dipolar coupling are hardly observed since the rapid tumbling of molecules averages them out, but in solid-state due to the immobilized states, they dominate the chemical shifts, which now with MAS can be removed (186).

Mimicking isotropic tumbling of the solution, ssNMR allows high-resolution measurements of large protein systems like amyloids, oligomers or membrane proteins, which would be difficult or impossible in solution (186). The strategy of assigning protein backbone sequences by ssNMR can be approached in two different ways, either by carbon or proton detection. Proton detection is more sensitive but requires higher spinning frequencies and are based on previously mentioned solution experiments, except they require a cross-polarization (CP) transfer step from proton to heteroatom to enhance heteroatom magnetization (187).

Besides, measuring large proteins, ssNMR has some advantages for relaxation studies too. The limit of measuring motions in solution is that internal motions are invisible due to them being slower than overall tumbling (nanosecond to microsecond). Since in ssNMR there are no internal tumbling in the sample, it is possible to record relaxations from picoseconds to milliseconds. This and the fact, the relaxations are mostly average in solution by inherent tumbling and in ssNMR is controlled by MAS, makes ssNMR a compelling technique to measure protein relaxations (188). On the whole, ssNMR is in a unique position, where it can be used from protein structural biology to cell walls and extracellular matrices (189).

Chapter III: Results Discussion and Summary

In this section, I will summarize the main results of my projects and go step by step on how I approached each of them. For full results, the reader is referred to the individual papers/manuscripts in the second half of the thesis.

3.1 The starting point of the projects (Paper V)

The PhD projects began with assisting in the purification of different chaperones for the α -synuclein study. The proteins were SurA, Skp, SecB, Trigger factor, which are found in *E. coli* periplasm/cytoplasm, mammalian Hsc70, and α -synuclein itself. The purified chaperones were used for the α -synuclein interactions titration by NMR and the BLI measurements, which resulted in the identifying the α -synuclein amino-terminal end as a binding site.

Proteins followed typical purification procedures of affinity column, followed by size-exclusion. SurA and Skp required extra refolding steps through dialysis to remove bound substrates. Hsc70 can have ATP or ADP bound states, which differ from the “empty” apo-conformation (190), therefore it also required extra optimization steps to remove bound states. Hsc70 was not stable during dialysis refolding, but on-column refolding worked. I was also familiarized with preparing NMR samples and setting up the experiments. The experience gained in this study was transferred directly to my main Ph.D. objective to study DegP serine protease as, for instance, the optimized Hsc70 purification protocol was adapted for DegP purification.

3.2 Probing the dynamics and interaction of DegP PDZ domains (Paper I)

The DegP is a periplasmic serine protease, which is activated in high temperatures. It can transform from hexamer to 24-meric substrate complexes and its oligomerization process is controlled by regulatory PDZ domains. My research goal was to determine the inherent dynamics of DegP, that control its transformation at different temperatures.

The first steps of characterizing DegP by NMR began with testing full-length proteolytic inactive DegP^{S210A}. The inactive form was chosen to avoid self-cleavage amino-terminal of DegP (97). The initial 2D [¹⁵N,¹H]-NMR spectrum of full-length DegP^{S210A} showed only a limited number of signals (~88 peaks from 448 expected). By testing a range of different temperatures (25°C, 37°C, 43°C, 50°C), I observed that DegP^{S210A} showed very high stability in higher temperatures and the number of peaks significantly increased at 50°C, at the same time being more uniformly distributed. However, the native slow tumbling of large proteins

leads to faster relaxation of transverse magnetization and lower signal intensities (191), therefore as DegP is a ~300 kDa hexamer, making it challenging to obtain good quality NMR data. To reduce DegP oligomer size, we exchanged tyrosine at the 444th position to alanine, as it had been identified to trimerize the protein (105). Albeit a little bit puzzling, since in the hexamer crystal structure of DegP^{S210A} (97), Y444 is oriented outside the structure and not interacting with any domain, whereas, in the 24-mer (93), it is a key residue keeping the cage complex together and is found in the PDZ1-PDZ2 interface (Fig. 9A). The DegP^{S210A,Y444A} provided a greatly improved 2D [¹⁵N,¹H]-NMR spectrum and we could continue with backbone assignments. This allowed us to assign 153 of 198 visible peaks and most of them were belonging to PDZ domains (Fig. 9B, C), accordingly, we decided to make an isolated PDZ1-PDZ2 construct and characterize it.

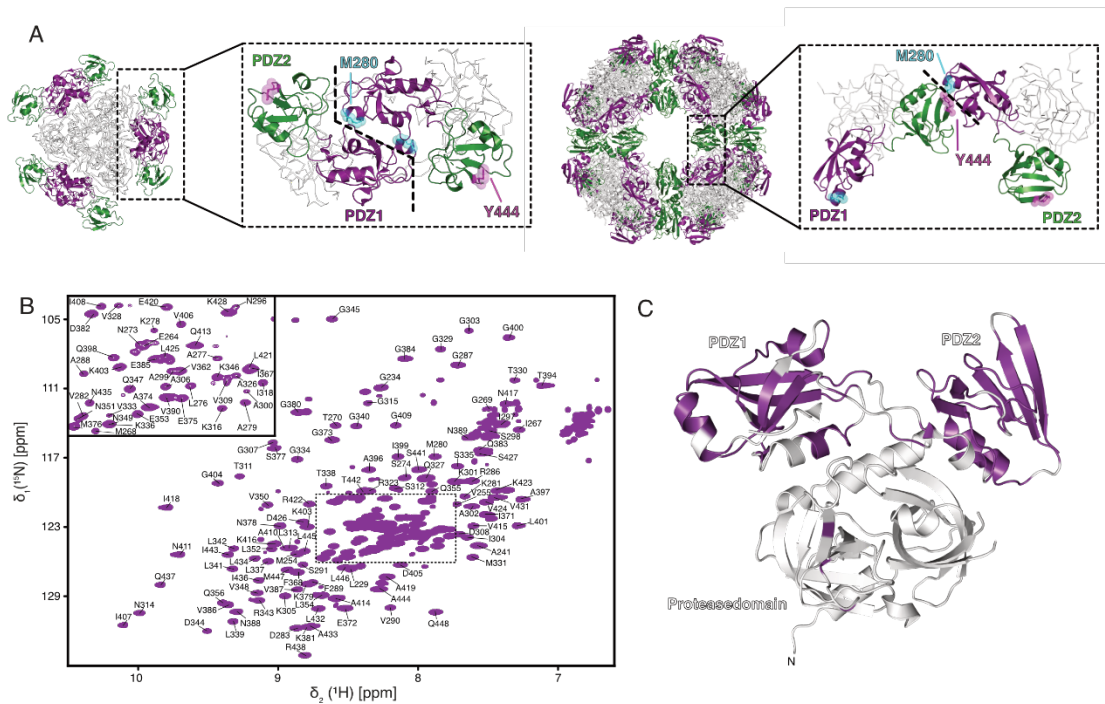


Figure 9. Full-length DegP^{S210A} in solution. (A) Comparison of the PDZ domains orientation between different monomers in the crystal structures of the hexamer (PDB-ID: 1KY9) and the 24-meric cage assembly (PDB-ID: 3OU0). Whereas in the inactive hexameric state this interaction is proposed to be mediated mainly via a direct interaction between two PDZ1s, the PDZ2s are found to be non-constrained at the hexamer edges, however, within the proteolytically active 24-mer the interaction between PDZ1 and PDZ2 of different molecules is achieved by a Y444 and M280 interaction indicating stabilization by sulfur- π -aromatic or a methyl- π -aromatic motif. (B) 2D [¹⁵N,¹H]-NMR spectrum of [^U-²H,¹⁵N,¹³C]-DegP^{S210A,Y444A}. The sequence-specific resonance assignment obtained from 3D TROSY-type triple-resonance experiments is indicated. (C) Cartoon representation of a DegP^{S210A} monomer from the crystal structure of the hexameric assembly of DegP^{S210A} (PDB-ID: 1KY9) with the sequence-specific resonance assignment indicated in purple. (B,C). Adapted and modified from from **Paper I**.

The PDZ1-PDZ2 construct displayed a remarkably good 2D [^{15}N , ^1H]-NMR spectrum at 50°C and allowed us to assign 150 out of a total of 188 PDZ1-PDZ2 resonances for backbone residues and their side-chains. As we were interested in DegP temperature transitions, we recorded a series of spectra at different temperatures. PDZ1-PDZ2 2D [^{15}N , ^1H]-NMR spectra showed significant differences from 25°C to 50°C, hence we needed to complete full assignments at 25°C to be able to trace residue chemical shifts. Comparing the secondary structure elements obtained from $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ correlations to random coil shifts did not show any significant deviations from PDZ1-PDZ2 structure within the hexamer crystal structure or between 25°C and 50°C chemical shifts. This let us conclude that our isolate PDZ1-PDZ2 domains form the same secondary structure as in the full-length DegP.

Advancing forward, we analysed various relaxation parameters (R_1 , R_2 , $R_{1\rho}$ hetNOE, S^2) at four distinct temperatures (25°C, 37°C, 43°C, 50°C). The average values of R_1 , R_2 were increasing/decreasing respectively with the temperature (Fig. 10A–C), indicating that the rotational correlation time τ_c was changing too. As τ_c is coupled to the molecular weight of the molecule, we considered that PDZ1-PDZ2 domains are undergoing oligomeric changes, but the τ_c values could be also attributed to the change of the viscosity. To address this, we compared theoretical τ_c values obtained from HYDRONMR relaxation prediction software (192) and the experimental τ_c values (Fig. 10D). The values in higher temperatures coincide, but deviated in lower temperatures, confirming our hypothesis, since HYDRONMR assumes PDZ1-PDZ2 domains as a monomer entity. We proposed that the PDZ1-PDZ2 domains are decoupled at 50°C, but interacting or oligomerizing at 25°C.

To explore the possible interactions of PDZ1-PDZ2, we titrated both isolated domains with each other and recorded a 2D [^{15}N , ^1H]-NMR spectra at each titration step (25°C, 37°C, 43°C, 50°C). It became apparent that the amino-terminus of the PDZ1 domain interacts with the carboxy-terminus of the PDZ2 domain. This was confirmed by the detailed analysis of chemical shifts perturbations and peak intensities. Furthermore, two residues were particularly affected: M280 and Y444 (Fig. 10E). Increasing the temperature appeared to weaken the interactions, suggesting that there is a specific temperature switch that disturbs the interface of PDZ1 and PDZ2 (Fig. 11E). We consider that it might be an oligomerization interface, thus we obtained translational diffusion coefficients *via* NMR diffusion experiment at a set temperature, which showed that DegP^{S210A} transitions from the hexamer to the trimer at 37°C–43°C, but the DegP^{S210A,Y444A} stayed as trimers (Fig. 10F). Additionally, size-exclusion elution traces DegP^{S210A} and mutant indicated hexamer and trimer elution profiles as well.

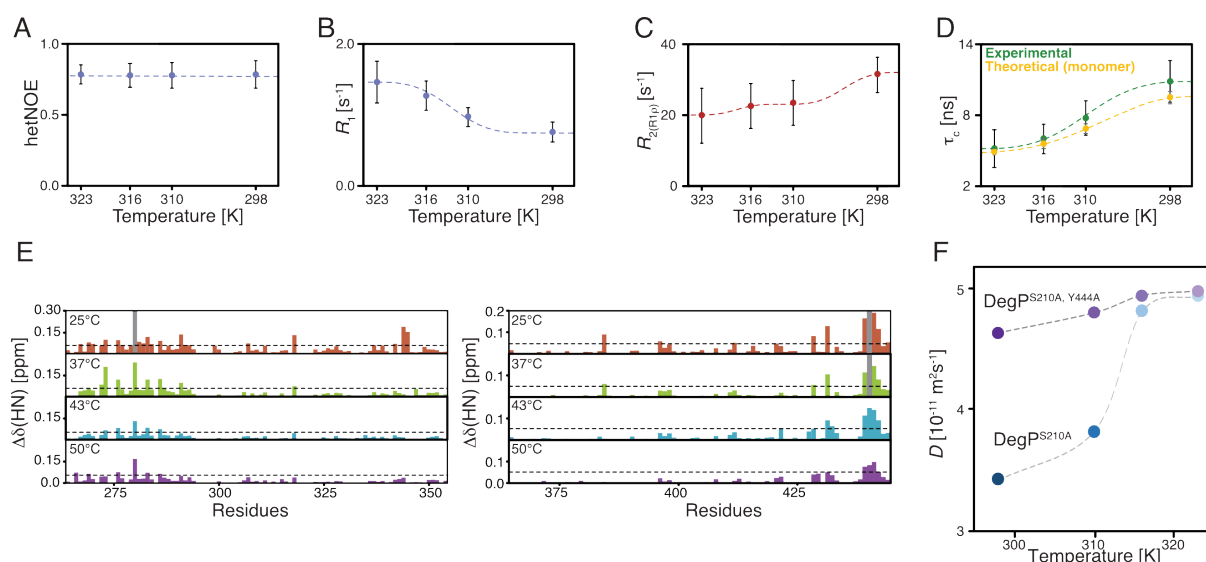


Figure 10. PDZ domains dynamics in different temperatures. Temperature dependence of the average hetNOE (A) R1 (B) and R2(R1ρ)-rates (C) over the indicated temperature range. Error bars are the S.D.. (D) Temperature-dependence of the measured rotational correlation time, τ_c , over the indicated temperature range (green). Theoretical values were calculated for the DegP crystal structure (PDB-ID: 3OU0) with HYDRONMR (72) for a PDZ1–PDZ2 monomer. (E) Detected chemical shift perturbations at four temperatures ranging from 25°C–50°C as indicated by the PDZ1 and PDZ2 domains. (F). Obtained molecular diffusion constants plotted against the temperature. The broken lines serve as a guide to the eyes only. Adapted and modified from **Paper I**.

As we were interested in Y444 and possible M280 interactions, we looked into their side-chain perturbations by the temperature. For Y444 residue, α were broadened beyond detection at 25°C and recovered intensity with an increase of temperature (Fig. 11A). Moreover, Y444 δ ring resonances were weakened as well at 25°C compared to 50°C. In comparison, Y444 ϵ signal intensities were stable over the range of the temperature (Fig. 11B). In opposition, M280 ϵ resonances showed slow chemical exchange, with two different states (Fig. 11C). As we contributed this to the M280-Y444 interaction, we defined them as open and lock states to describe broken or existing contact. Close inspection of available crystal structures revealed the presence of the same interaction of M280-Y444 in the Deg 24-mer cage (100). To test a possible stabilizing methionine–phenylalanine interaction motif (193), the two mutants DegP^{S210A, M280A}, DegP^{S210A, M280I} and their respective isolated PDZ domains were made and tested for diffusion experiment or titrations. All DegP mutants were identified as trimers *via* diffusions experiments and their PDZ domains did not interact with each other, except for PDZ1^{M280I}, presumably due to longer side chain and methyl/tyrosine π -orbitals interactions (194).

To complement our NMR results with biochemical data, we performed cleavage experiments of β -casein, using fluorescent reporter peptide and SDS-PAGE analysis. Proteolytic assays

indicated that all mutants cleave faster than wild-type DegP (Fig. 11E, F). Moreover, mixtures of β -casein and each DegP variant were loaded to the size-exclusion column. β -casein cage complex formations were not formed by DegP^{S210A,M280A} and DegP^{S210A,Y444A}. However, DegP^{S210A,M280I} was still able to form cages but that was expected due to methyl/tyrosine π -orbitals interactions.

In conclusion, NMR spectroscopy enables us to capture the transitions of DegP from hexameric to trimeric state at a higher temperature that is governed by PDZ1-PDZ2 domains methionine-tyrosine lock. We proposed a DegP activation mechanism, where DegP can be activated the substrate binding in the lower temperatures, but induced by temperature stress, DegP activates

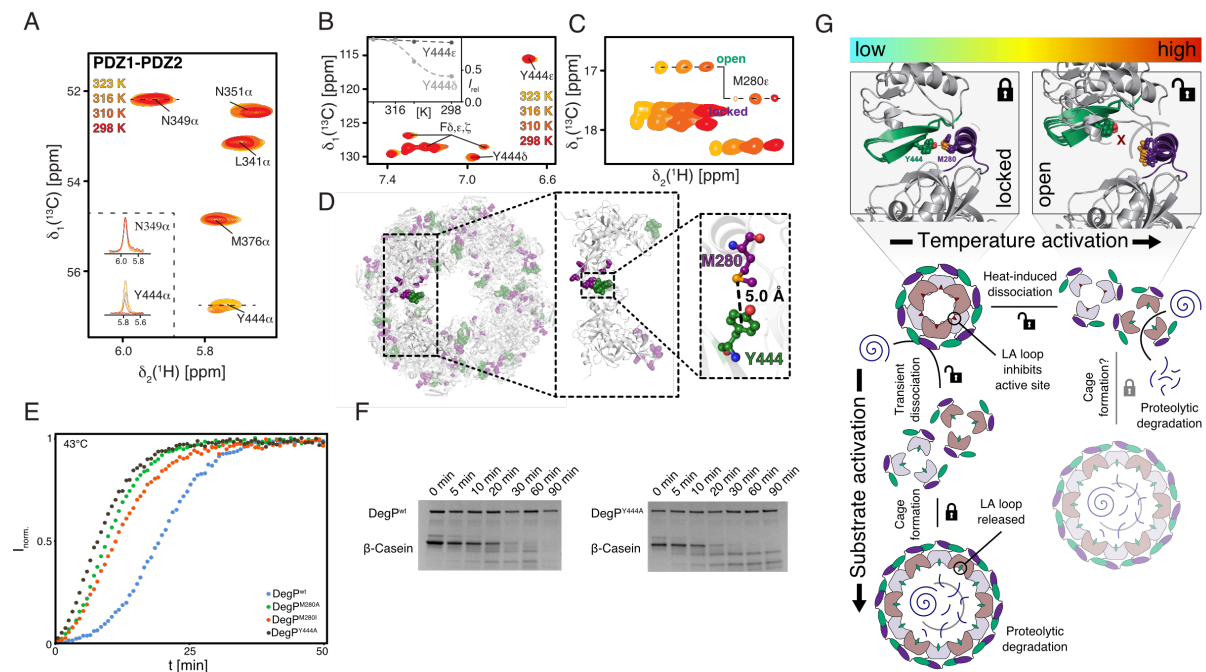


Figure 11. (A) Focus on the C α resonance region around Y444 of a 2D $^{13}\text{C}, ^1\text{H}$ -NMR spectrum of [U- $^{13}\text{C}, ^{15}\text{N}$]-PDZ1-PDZ2 at 25°C–50°C as indicated. The inset shows the respective $\delta_2(^1\text{H})$ -1D cross-sections through the N349 α and Y444 α resonances. (B) Aromatic region of a 2D $^{13}\text{C}, ^1\text{H}$ -NMR spectrum of [U- $^{13}\text{C}, ^{15}\text{N}$]-PDZ1-PDZ2 at 25°C–50°C as indicated. Inset shows the signal intensities of the aromatic Y444 ϵ and Y444 δ resonances in the dependence of the temperature. (C) Methionine- ϵ -methyl region of a 2D $^{13}\text{C}, ^1\text{H}$ -NMR spectrum of [U- $^{13}\text{C}, ^{15}\text{N}$]-PDZ1-PDZ2 at 25°C–50°C as indicated. Spectra were manually shifted along the ^1H dimension to illustrate the transition between open and locked states. Residues, CSPs twice the S.D. on PDZ1, are highlighted in purple (G266, G269, K278, M280, D283, and R286), whereas residues in PDZ2 are highlighted in green (A433, N435, I443, Y444, and L445). The enlargement focuses on the central M280–Y444 interaction at one of the inter-domain interfaces. (E) DegP (10 μM monomer concentration) proteolysis of β -casein (50 μM) via detection by a fluorescent reporter peptide (100 μM) at 43°C. Experiments were performed as triplicates yielding an identical result. (F) SDS-PAGE analysis of β -casein (5 μM) cleavage by DegP and DegP^{Y444A} (1 μM monomer concentration) at the indicated time intervals at 25°C. (G) Proposed temperature activation of DegP mechanism. Adapted and modified from **Paper I**.

itself and disassociates to trimeric pieces (Fig. 11E). Moreover, a new study by the Lewis E. Kay group proposed two DegP oligomerization pathways based on the dynamic light scattering experiments, suggesting that DegP hexamers at lower temperatures ($\geq 25^{\circ}\text{C}$) are kept together by the presumed PDZ1-PDZ1 interactions (as observed in the hexameric crystal structure) and at higher temperatures by PDZ1-PDZ2 domains, which the authors could also show by NMR experiments as well (195). This matches with our findings and raises the question of whether the established inactive hexamer crystal structure is only one of several possible conformations.

3.3 Investigating DegP protease core (Paper II)

After we finished our study about DegP PDZ domains, we continued with the rest of the DegP, by starting the analysis of the DegP protease core. We were interested in catalytic (L1-3) and allosteric loops (LA, LD), that are hidden inside hexameric core (Fig. 12A). The LA loop connects two DegP monomers (Fig. 12B), whereas interplay L1-3 and LD help form an active centre (Fig. 13C).

We truncated DegP^{S210A} construct to the first 259 residues (DegP-NTD^{S210A}), which included the full sequence of protease domain and prepared a sample for NMR measurement. The first 2D [¹⁵N, ¹H]-NMR spectrum of DegP-NTD^{S210A} showed 143 peaks from 279 expected, which were too few to continue. Compared to the related HtrA family members, we recognized that the DegP protease core is very similar to one of the DegP human homologs – HtrA2. HtrA2 is observed as a trimer, but a mutation in the protease core (F149D), monomerizes it (80). F149D is located in the centre of the HtrA2 protease interface. The structural alignment of HtrA2 with DegP matched HtrA2s F149 position with DegPs Y163 residue. DegP-NTD^{Y163D, S210A} construct led to a highly improved [¹⁵N, ¹H]-NMR spectrum, due to monomerization, with almost all of the expected peaks. After running and analysing standard backbone experiments, we were able to assign 222 peaks from 274 observed. The secondary structure based on the C $^{\alpha}$, C $^{\beta}$ chemical shifts, indicated a good agreement between crystal structure and the secondary structure elements in solution (Fig. 12D).

As initially mentioned, we were intrigued by inherent loops dynamics, since some of them (LA, L3) are not very well resolved in available crystal structures and are only available in the modelled structure (196). We used a combination of PRE and relaxation measurements to determine the dynamics of loops. For PRE, we mutated two cysteines (C57S or C69S) in the LA loop of DegP-NTD^{Y163D, S210A} and chemically labelled only one at the time with MTSL reagent. Due to the free electron of the reagent, it will broaden all resonances, which are

spaciously close to cysteines. The PRE results suggested that LA motions are unexpectedly strong, reaching the L3 loop, which is oriented farther away than L1 and L2 loops in DegP hexameric structure (Fig. 12C).

To extract meaningful relaxation data, we used model-free analysis, which takes into account all recorded relaxation R_1 , R_2 , hetNOE parameters and an available structure model. As expected, loops had increased flexibility compared to folded parts of the protease. However, the T176 and S134, which are not part of the mentioned loops, had also increased relaxation rates (Fig. 12F). A closer inspection of available structures revealed that T176 formed hydrogen

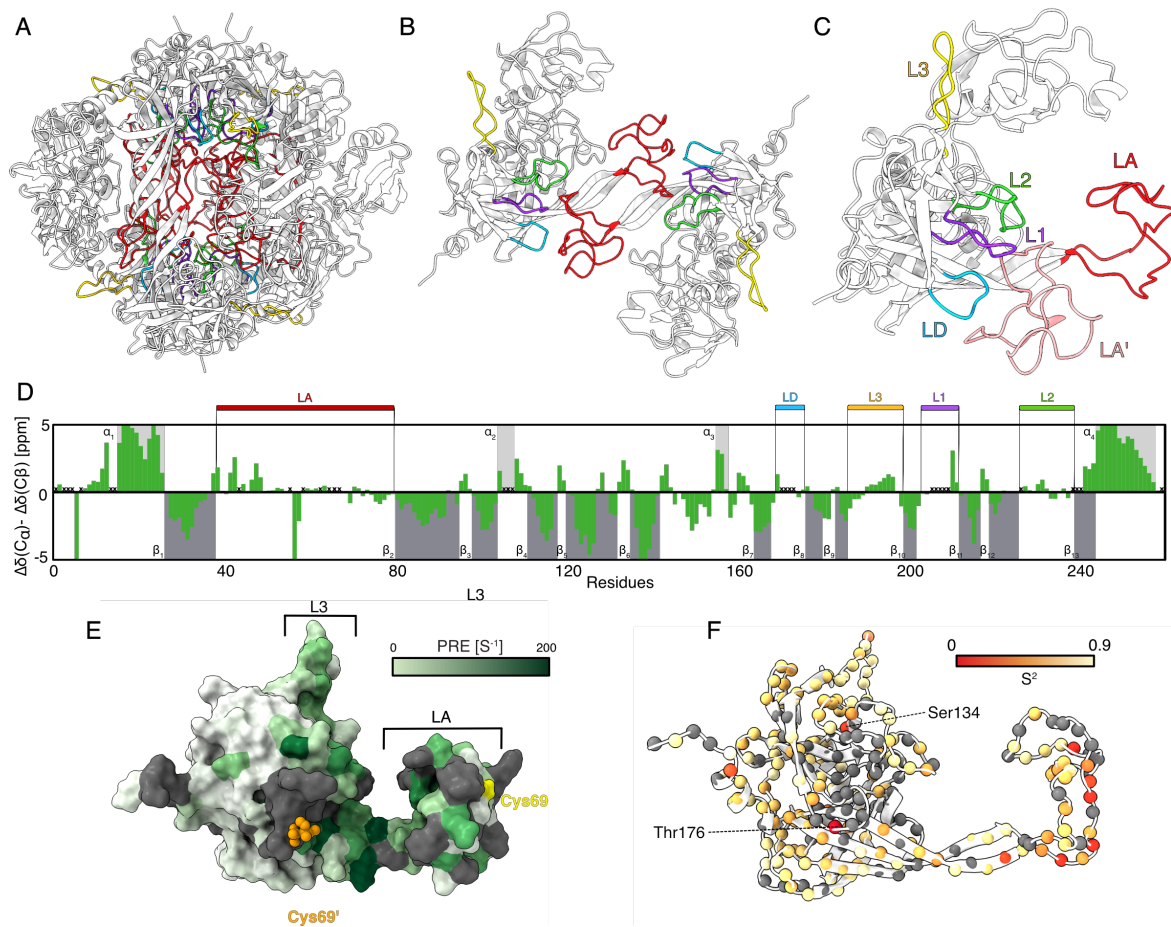


Figure 12. DegP protease domain architecture. DegP catalytic and allosteric loops as seen in hexameric structure (A), on focused dimers (B) and monomer (C) (Loops colored: L1- purple, L2-green, L3-yellow, LA red, LA' (from adjacent monomer)- light red). (D) Secondary backbone ^{13}C chemical shift analysis plotted against the DegP-NTD^{Y163D,S210A} residues. X symbols present unassigned residues. The secondary structure elements derived from the crystal structure with computationally modelled loops (PDB-ID: 1KY9) are highlighted in grey. (E) The PRE rates plotted against DegP-NTD^{Y163D,S210A} surface structure. The rates values are indicated by the white to dark green gradient. Grey color shows missing values. (F) Local backbone dynamics on sub-nanosecond timescale order parameter S^2 at 37°C. The amide moieties of the DegP-NTD^{Y163D,S210A} construct are shown as spheres and the hetNOE values are indicated by the white to blue gradient. Adapted and modified from **Paper II**.

bonds with R187 and Q200 in DegP 24-mer (Fig, 13A), but not in the DegP hexamer. Our data suggest that this bond could govern T176 inherent motions during allosteric activation since it has been observed that the T176V mutation abolishes the proteolytic activity of DegP (197). Contrarily S134 does not form any bonds in the DegP 24-mer or DegP hexamer, but its adjacent residue R133 forms hydrogen bonds with E193 and N194 residues of the L3 loop (Fig. 13B). As the L3 loop is not well-resolved in the hexameric structure, we speculate that the R133 or S134 together with L3 might work as another allosteric site. To confirm this hypothesis, we will need to test the proteolytic activity of DegP^{R133V} and DegP^{S134V} mutants.

Taken together, for the first time DegP protease domain was assigned by NMR spectroscopy and new allosteric dynamics were observed in the core residues. Moreover, the strong LA motions could be observed, which potentially could act as gates, for opening and closing the active site to the substrate. For completeness, we aim in including relaxation data of side-chain methyl groups of the protease domain (DegP-NTD^{Y163D,S210A}). The data had been already recorded and we are currently completing the side-chain assignments to enable analysis.

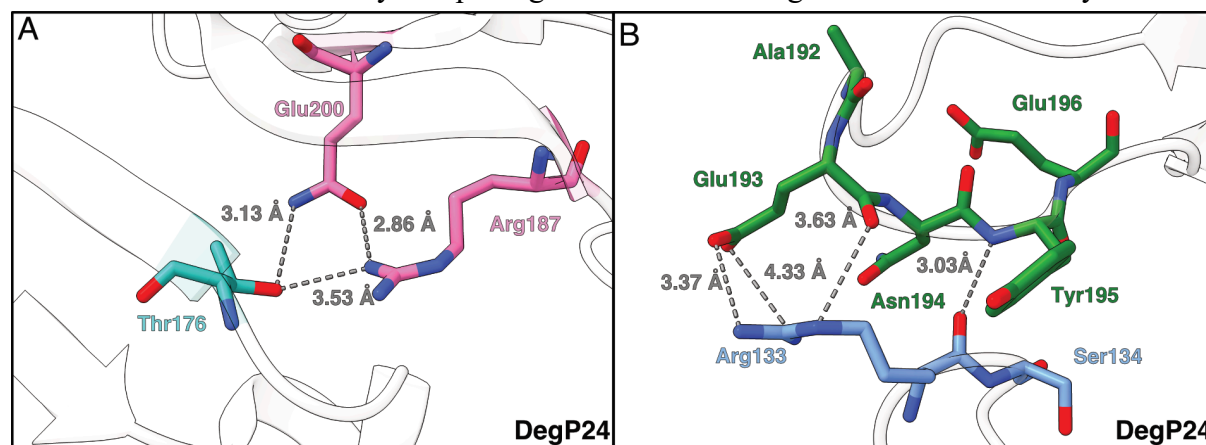


Figure 13. DegP allosteric regulation. Hydrogen network bonds between Thr176 and adjacent monomer Arg187, Glu200 in DegP 24-mer (A). The Arg133 hydrogens bonds to L3 residues Glu193, Asn194 (B). (PDB id: 3OU0) Adapted and modified from **Paper II**.

3.4 Insight to full-length DegP (Paper III)

In spite of establishing dynamics and interactions of the individual domains of DegP, our final ambition was to capture full-length DegP dynamics by NMR. Our solution 2D [¹⁵N,¹H]-NMR spectrum of DegP^{S210A} had very few signals even at 50°C, so we decided to apply ssNMR as it is more suited for large protein complexes (198). In collaboration with Paul Schanda (Austria), who is an expert in solid-state NMR, we recorded an initial 2D sNH spectrum, that showed 172 peaks out of the 448 expected. As the quality and sensitivity of the sample were sufficient, we continued the backbone assignments experiments (hCONH, hCANH, hcaCBcaNH, hCOcaNH,

hCAcoNH, hNhhNH). By the time of writing this thesis, not all experiments are recorded yet and current backbone assignments lack the critical resolution for assignments. However, 21 peaks could already be assigned by superimposing solution 2D spectra of individual DegP domains [$^{15}\text{N}, ^1\text{H}$]-PDZ1-PDZ2 or 2D [$^{15}\text{N}, ^1\text{H}$]-DegPNTD $^{\text{Y163D}, \text{S210A}}$) with solid-state 2D spectrum of [$^{15}\text{N}, ^1\text{H}$]-DegP $^{\text{S210A}}$ and confirmed by solid-state 3D experiments (Fig. 14A,B).

We are currently recording additional 4D experiments (starting from hCOCANH), which should help us greatly with the assignments. Once we assigned the full-length inactive hexamer, we envisage determining relaxation parameters and potentially observing DegP cages complexes by ssNMR in presence of oligomerizing peptides, which would complementary data to previous x-ray crystallography and cryo-EM structures of DegP cages (93, 100).

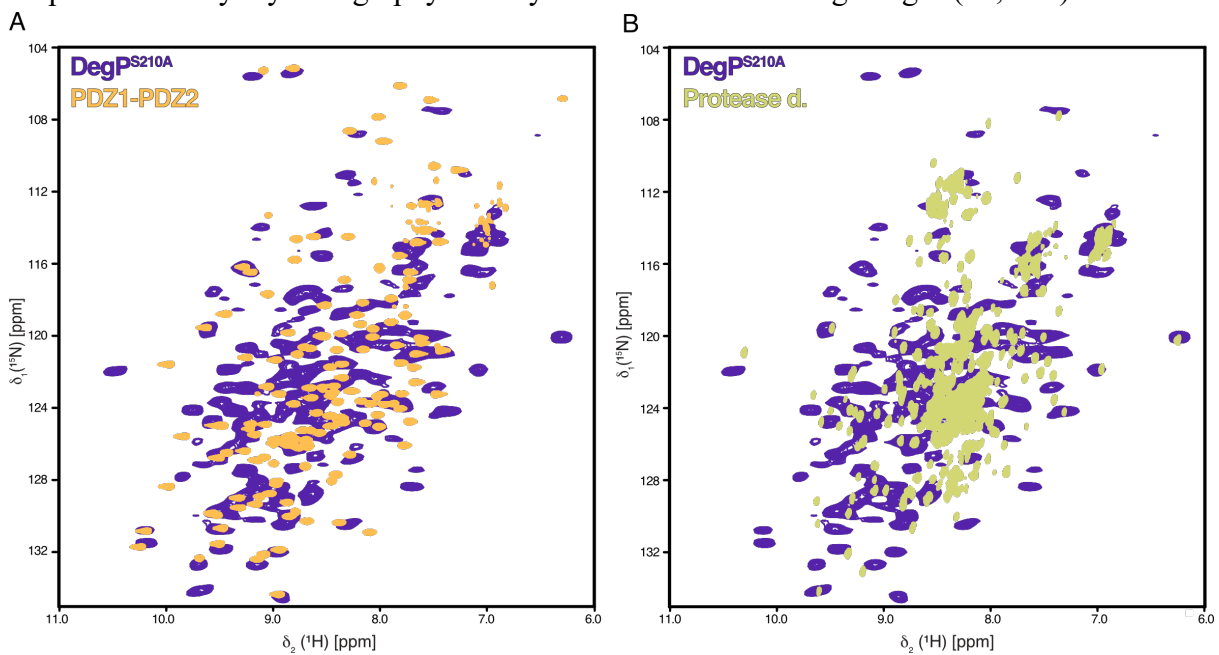


Figure 14. Overlay of DegP and its individual domains. (A) 2D [$^{15}\text{N}, ^1\text{H}$]-ssNMR spectra of [$U\text{-}^{13}\text{C}, ^{15}\text{N}, ^2\text{H}$]-DegP $^{\text{S210A}}$ (A, B, purple) and 2D [$^{15}\text{N}, ^1\text{H}$]-NMR spectra of [$U\text{-}^{15}\text{N}$]-PDZ-PDZ2 (A, orange) and [$U\text{-}^{15}\text{N}, ^2\text{H}$]-DegPNTD $^{\text{Y163D}, \text{S210A}}$ (B, green) Adapted and modified from Paper III.

3.5 Interaction between Skp and DegQ/DegP proteases (Paper IV)

One of the main objectives of this thesis was to study the potential interplay between the DegP serine protease and the Skp chaperone since they are regulated by the same σ^{E} and Cpx regulons (Fig. 3) (199, 200). Moreover, they function in parallel at the secondary pathway for transporting OMPs in the periplasm (38). Although there were attempts to deduce DegP and Skp direct interactions, they have not been successful so far (46, 201). However, it has been recently proposed that Skp might be in an unfolded-monomer structure in the periplasm and that it only forms stable trimers when it is bound to the client (12). In addition, DegP has been

shown to cleave unfolded proteins and to degrade CpxP, another periplasmic chaperone (202). Here, we wanted to investigate the Skp conformation inside the cell and explore the possibility of the DegP acting as a “housekeeper” by removing residual monomerized Skp.

First, we analysed the protein conformation in solution purified Skp, monomerized mutant (Skp^{A108R}) and a Skp in outer membrane vesicles (OMVs), that latter reflecting Skp state in the native environment. Uniformly distributed resonances in 2D [¹⁵N,¹H]-NMR spectra of a purified Skp indicated that it was folded. On contrary, Skp^{A108R} and Skp in OMVs showed a different conformation, as most of the peaks were distributed around 8 ppm, indicating proteins are partially unfolded (Fig. 15).

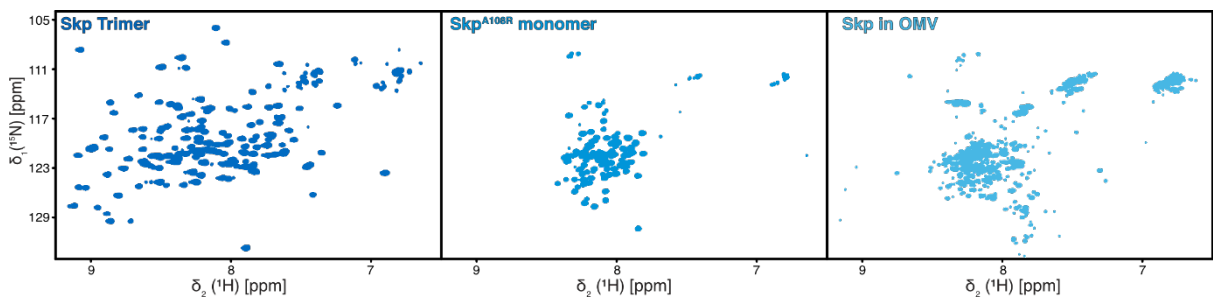


Figure 15. Different Skp conformations in-solution. 2D [¹⁵N,¹H]-NMR spectra of [^U-¹⁵N,²H]-Skp 100 μM, [^U-¹⁵N,²H]-Skp^{A108R} 50 μM and [^U-¹⁵N,²H]-Skp in OMVs. Adapted and modified from **Paper IV**.

In the next step, we tested the direct interactions of Skp with DegP. NMR Titration experiments between Skp and DegP^{S210A} revealed a slight line broadening of the resonances, suggesting very weak or inexistent interaction (Fig. 16A). On the other hand, in Skp^{A108R} and DegP titration, we observed stronger line shape broadening (Fig. 16B). We decided to test DegQ

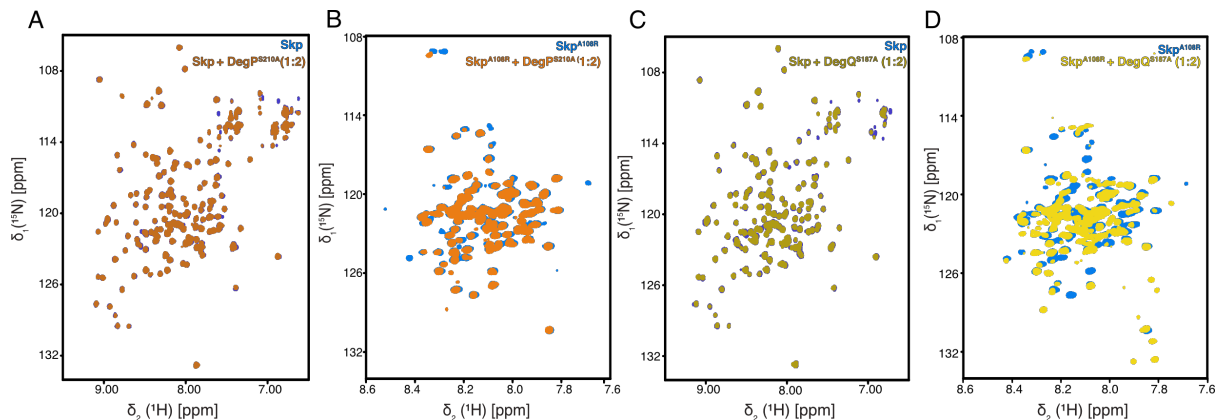


Figure 16. DegP and DegQ display increased interaction with Skp monomer. 2D [¹⁵N,¹H]-NMR spectra of [^U-¹⁵N,²H]-Skp 100 μM (A, C, blue) and [^U-¹⁵N,²H]-Skp^{A108R} 50 μM (B, D, blue) as well as after the addition of two molar equivalents of the DegP (orange) or DegQ (yellow) respectively. Titrations were measured in 25 mM MES, 150 mM NaCl pH 6.5 at 37°C. Adapted and modified from **Paper IV**.

protease as well, which is the closest HtrA family member to DegP and should have similar substrate recognition mechanisms. DegQ, as well, did not indicate any interaction with trimer Skp (Fig. 16C), but for Skp^{A108R} showed a very strong binding, as resonances were broadened beyond detection with new signals appearing (Fig. 16D). Taken together, we concluded that Deg proteases explicitly recognized the monomerized Skp with DegQ having a stronger affinity than DegP.

As we confirmed interaction, we tested cleavage of Skp variants by DegP and DegQ *via* fluorescent assay using fluorescent reporter peptide (100) ((Fig. 15). We included an additional mutant Skp^{A108L}, which has a trimer conformation at 25°C and a monomer at 37°C (12). The trimeric Skp was not degraded by DegP or DegQ, whereas Skp^{A108R} was cleaved by both at 25°C and 37°C. The semi-stable mutant Skp^{A108L} was cleaved at 25°C and 37°C by DegQ, but not by DegP at 25°C, most likely due to DegQ higher affinity to Skp monomers, that could exist in semi-stable confirmation. We confirmed proteolysis of Skp^{A108L} SDS-PAGE, and send the lower band of cleaved Skp^{A108L} for the mass spectrometric analysis, where it was confirmed that the first 28 residues (from the canonical sequence) were cleaved. Interestingly, these residues form a crucial β -strand in the Skp crystal structure and a part of the trimerization interface (Fig. 18).

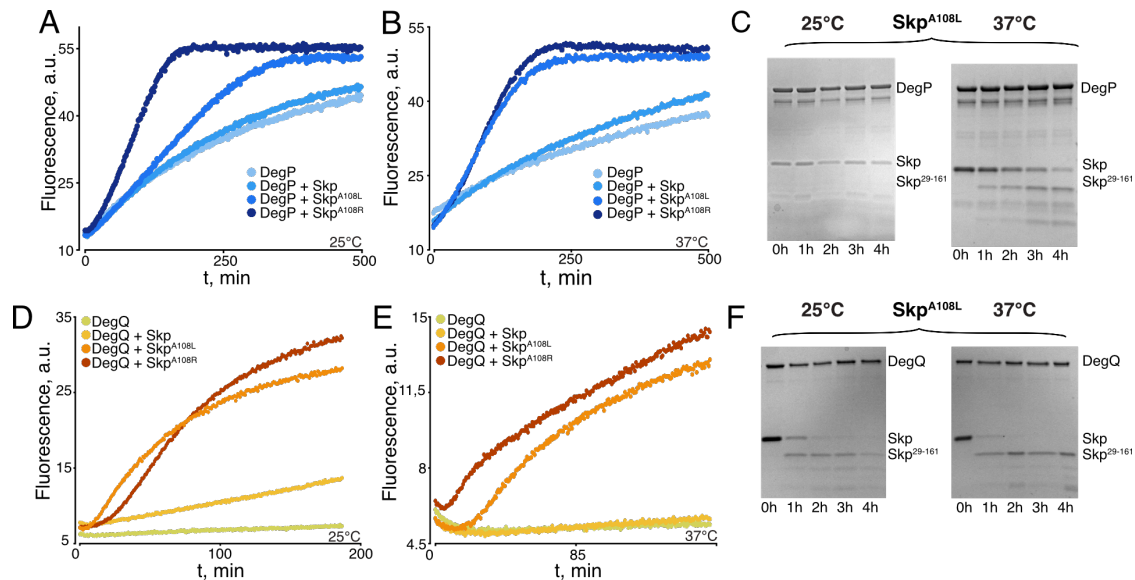


Figure 17. Proteolysis of Skp and its variants by DegQ or DegP (A, B) DegP and (C, D) DegQ (10 μ M monomer concentration) proteolysis of Skp, Skp^{A108L}, Skp^{A108R} (50 μ M) via detection by a fluorescent non-activating reporter peptide (100 μ M) at 25°C (A, C) and 37°C (B, D). (E, F) SDS-PAGE analysis of Skp^{A108L} (5 μ M) cleavage by DegP (5 μ M monomer concentration) or DegQ (1 μ M monomer concentration) at the indicated time intervals at 25°C and 37°C. Fluorescence experiments were performed as triplicates yielding identical results and representatives are shown. Adapted and modified from **Paper IV.**

All in all, we confirmed Skp monomer proteolysis by DegP and DegQ, but we are looking at how it connects to the *in vivo* processes. Our aim is to observe Skp expression and proteolysis in wild type MG1655 strain and in *degP degQ* knockouts, which should help connect our observations in *in vitro*. We believe our final results will be another missing piece of the puzzle in understanding Skp regulation and Deg proteases function and their relationship in *E. coli* periplasmic protein quality control.

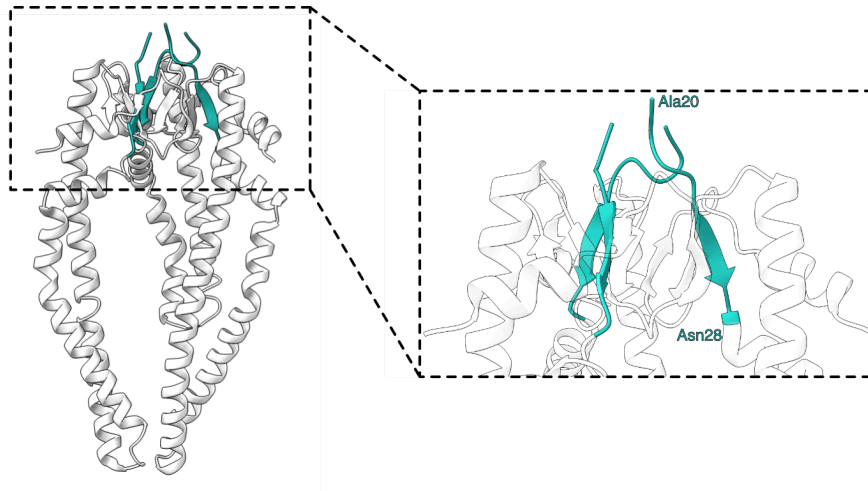


Figure 18. The Skp^{A108R} cleavage site. The DegQ and DegP cleave the first eight residues (not including histidine tag or signal sequence) of the Skp^{A108R}, which corresponds to the trimerization interface in Skp (PDB id:1SG2). Adapted and modified from **Paper IV**.

3.6 Concluding remarks.

During the Ph.D. project, I was able for the first time to characterize DegP by NMR spectroscopy. I observed the DegP oligomerization patterns in different temperatures and indicated PDZ domains as the main source of it. I and my co-workers established that DegP trimers in solution are connected through PDZ1 and PDZ2 domains by methionine-tyrosine motif, in contrast to previously seen PDZ1 and PDZ1 trimer interface in the crystal structure (203). The existence of possible two DegP hexamers populations and that hexamers discharge to trimers in high temperature shows the DegP has strong self-regulating properties that coordinate its activity. In tandem with PDZ domains, protease core loops act as a second self-control apparatus. The proteases domain inherent dynamics of allosteric residues and LA inhibitory loop ultimately might be a key determining the protease activity. Furthermore, the transferred knowledge of individual DegP domains by in solution NMR to full-length DegP by solid-state NMR will help further develop an understanding of DegP protein quality control capabilities in the periplasm. In addition, the Skp-monomer proteolysis by DegP and DegQ serine proteases revealed their new functions in an overreaching network of envelope chaperones and proteases.

On the whole, each project nicely led to the next one (Fig. 19) and established the basis for the HtrA protein analysis by NMR spectroscopy. It will be intriguing to see if the rest members of the HtrA family follow the dynamic properties that I observe in DegP:

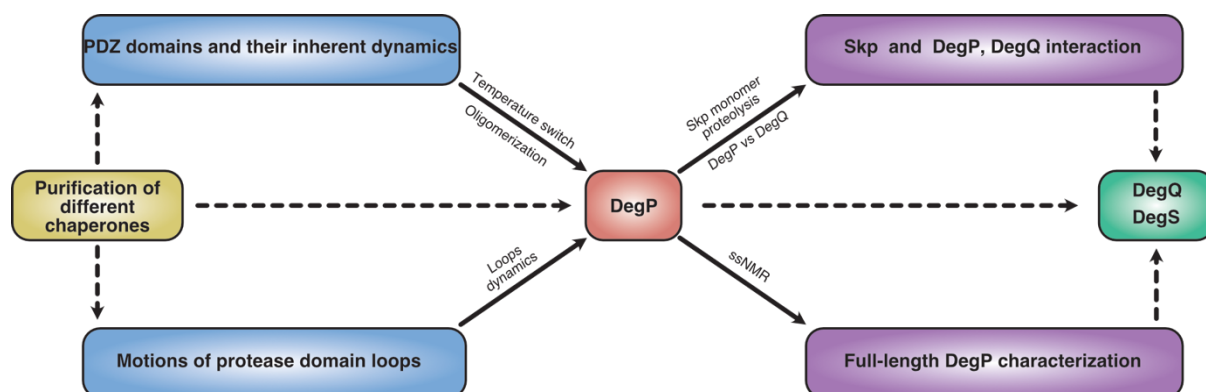


Figure 19. The flow chart of the thesis projects. The initial lessons from the purification of different chaperones aided in starting DegP projects (protease and PDZ domains). After obtaining data on the individual DegP domains, we transferred this information to full-length DegP analysis and interaction study of Skp with DegP/DegQ. Finally, the overarching goals are to use DegP as starting point for other HtrA proteases (DegQ, DegS) studies by NMR spectroscopy.

Chapter IV: Future perspectives

The characterization of DegP in this thesis study revealed novel functional DegP elements, that can be exploited for further analysis. The first time almost complete backbone assignments of DegP protease and PDZ domains, together with their dynamics properties can be used for further interactions investigations or even antibiotic developments. Moreover, there are still additional questions to investigate if DegP can form cage-like structures in the cell (due to physical size limitations) or intermediate steps of DegP activation. Furthermore, as we observed Skp in a vastly different state from what was established, there is a possibility other periplasmic chaperones like Spy, FkpA, SurA might have different conformations in the envelope that make them susceptible to proteases. Finally, we already decided to apply the learned experiences from DegP to other periplasmic proteases DegS and DegQ. The early observations of their PDZ domains, already indicate that they are excellent systems to analyze by NMR spectroscopy due to well disperse resonances and stable fold (Fig. 20). Interestingly, a recent report about human HtrA2 (204), hints that our observed PDZ domains interaction could be prevalent not only to bacterial HtrAs and might be a conserved interaction in the HtrA family.

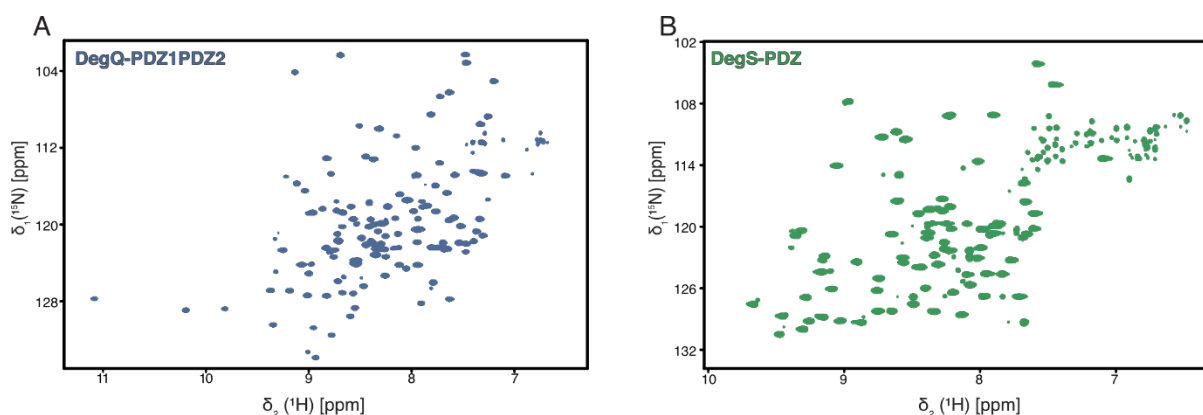


Figure 20. DegQ and DegS PDZ domains in-solution. 2D [^{15}N , ^1H]-NMR spectra of [U - ^{15}N]-DegQ-PDZ1PDZ2 1mM (A) and [U - ^{15}N]- DegQ-PDZ 1mM (B) measured in 25 mM 25 mM K-pi, 1 mM EDTA pH 7.0 at 37°C.

Acknowledgements

In the last part of the thesis, I want to thank all the people, who helped me during the last four years. It will be less formal and prepare for a lot of “thank you” or “good luck” (and I know some of you will skip the entire thesis, just to read this part, so I tried to mention most of you).

First and foremost, I bow down to my supervisor **Björn** for giving me a chance to do a Ph.D. in his group and for the enormous amount of work he put in helping me. Your determination and commitment to science is a thing to admire.

I thank **Yosh** for all the scientific and sometimes (when it was needed) not-so-scientific advices as well as for supporting my work. Also, I can not thank enough **Lisa** for giving more than the required time in teaching molecular biology skills and reading my messy writings.

Ashish, I appreciate that you always find time to explain NMR and daily help. Likewise, I am grateful to **Irena** for lab assistance and for sharing homesickness thoughts. And for **Laura**, who might not see this, thanks for the support, even from France, and for being the funniest person in the lab.

I wish all the best and luck to **Emelie** in motherhood and finishing up the PhD. **Damasus**, good luck in completing your last steps, you can do it. **Ylber**, thanks for all fun and being the soul of doctoral students, the driving force of all research, a pillar that supports Lundberg, **Jens**, starting now, you will be in all acknowledgements for all scripts you have developed for us, thank you. For our newcomers, **Filippo** and **Hannah**, good luck in starting up your projects and do not forget those shovels for the trip.

For the first floor (dungeon) colleagues (**Johanna**, **Dimitra**, **Lisa**, **Katharina**, **Vajradhar**). Although we did not have much time to interact I wish you all the luck in cryo-EM studies (and, yes, **Dimitra**, I was jealous of all the food you brought!)

Moving to the second floor, I want to thank **Kristina**, for always being calm and supportive during the KEM350 course, even when we complained about half of the course students. **Jessica**, thank you for your everyday cheerful greeting and congratulations to **Florian**, on his “postdoc” career, I miss our geeky discussions.

Thank you to my examiner, **Richard**. I will never forget your political questions about Trump and Putin during the first interview, that was not what I expected coming here, but it helped me to relax.

Andrea quit smoking! But more importantly, relax and have fun. **Ann**, thank you for being a faithful badminton player, it was challenging but fun to play with you. **Amke**, I hope you are enjoying Copenhagen and good luck in wrapping up your studies.

Andreas, your Swedish spirit is unmatched and thank you for taking care of all the course lab management. **Per**, probably expected, but thanks for watching over of the beer club and good luck in writing the thesis.

Lucija, thank you for joining the äkta club and not shrugging away. **Leona**, **Analía** thank you for enduring my grumpiness during Äkta help.

Jonatan, thank you for keeping the team together, even when I wanted to run away from you. **Owens**, thank you for being light-hearted and always in good spirits. **Arpitha**, you are already proficient in many areas, so getting a driving license will be effortless. **Adams**, it was nice to know there is another food truck guy lover here (fingers crossed for him improving back to his original standards). **Dylan**, godspeed on your Ph.D journey.

Good luck **Viktor**, **Maja**, **Daniel**, **Greger**, **Giorgia** in writing their thesis too, hopefully, it will be a breeze. **Lidija**, **Doris**, **Laras** good luck in your second half, it goes faster than you think.

Tinna, wishing you lots of luck in parenthood! **Weixiao**, all the best in your career.

Thank you to all remaining researchers and PIs of Lundberg (**Sebastien**, **Julia**, **Gergely**, **Gisela**, **Michal**, **Swagatha**) for creating an academic environment and supporting students.

Also, shout-out to previous doctoral students (**Elin**, **Emil**, **Petra**, **Majo**, **Rajiv**, **Cecilia**, **Léo**) who gave all initial introductions to the Lab and made Lundberg an exciting place to work.

Gratitudes to NMR Swedish Center staff (**Göran**, **Vladislav**, **Ulrika**, **Daniel**, **Anders**, **Arthur**, **Zoltán**) for always quick help with instruments and keeping up with all the bookings.

Many thanks to my mother, **Vilija**, and sister, **Vilma**, for supporting my choices and calling every week to discuss Lithuanian and Swedish news.

I hope some of you will come to visit me in Lithuania and in return I will be always able to be a guest in Sweden.

References

1. D. A. White, A. K. Buell, T. P. J. Knowles, M. E. Welland, C. M. Dobson, Protein Aggregation in Crowded Environments. *J. Am. Chem. Soc.* **132**, 5170–5175 (2010).
2. A. C. Miklos, M. Sarkar, Y. Wang, G. J. Pielak, Protein crowding tunes protein stability. *J. Am. Chem. Soc.* **133**, 7116–7120 (2011).
3. Y. E. Kim, M. S. Hipp, A. Bracher, M. Hayer-Hartl, F. Ulrich Hartl, *Molecular Chaperone Functions in Protein Folding and Proteostasis* (2013; <http://www.annualreviews.org/doi/10.1146/annurev-biochem-060208-092442>), vol. 82.
4. B. Chen, M. Retzlaff, T. Roos, J. Frydman, Cellular strategies of protein quality control. *Cold Spring Harb. Perspect. Biol.* **3**, 1–14 (2011).
5. J. P. Hendrick, F. Hartl, Functions of Heat-Shock. *Annu. Rev. Biochem.* **62**, 349–84 (1993).
6. C. M. S. Kumar, S. C. Mande, G. Mahajan, Multiple chaperonins in bacteria—novel functions and non-canonical behaviors. *Cell Stress Chaperones.* **20**, 555–574 (2015).
7. D. K. Clare, H. R. Saibil, ATP-driven molecular chaperone machines. *Biopolymers.* **99**, 846–859 (2013).
8. D. Huber, N. Rajagopalan, S. Preissler, M. A. Rocco, F. Merz, G. Kramer, B. Bukau, SecA Interacts with Ribosomes in Order to Facilitate Posttranslational Translocation in Bacteria. *Mol. Cell.* **41**, 343–353 (2011).
9. W. Voth, U. Jakob, Stress-Activated Chaperones: A First Line of Defense. *Trends Biochem. Sci.* **42**, 899–913 (2017).
10. O. Suss, D. Reichmann, Protein plasticity underlines activation and function of ATP-independent chaperones. *Front. Mol. Biosci.* **2**, 1–10 (2015).
11. U. Jakob, W. Muse, M. Eser, J. C. A. Bardwell, A. Arbor, Chaperone Activity with a Redox Switch. **96**, 341–352 (1999).
12. G. Mas, B. M. Burmann, T. Sharpe, B. Claudi, D. Bumann, S. Hiller, Regulation of chaperone function by coupled folding and oligomerization. *Sci. Adv.* **6** (2020),

doi:10.1126/sciadv.abc5822.

13. O. P. Ward, in *Comprehensive Biotechnology* (Elsevier, 2011; <https://linkinghub.elsevier.com/retrieve/pii/B9780080885049002221>), pp. 571–582.
14. K. M. Heutinck, I. J. M. ten Berge, C. E. Hack, J. Hamann, A. T. Rowshani, Serine proteases of the human immune system in health and disease. *Mol. Immunol.* **47**, 1943–1955 (2010).
15. R. Rosen, D. Biran, E. Gur, Protein aggregation in Escherichia coli : role of proteases. **207**, 9–12 (2002).
16. A. R. Khan, M. N. G. James, Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci.* **7**, 815–836 (1998).
17. C. Lazure, The Peptidase Zymogen Proregions: Natures Way of Preventing Undesired Activation and Proteolysis. *Curr. Pharm. Des.* **8**, 511–531 (2005).
18. A. Chakrabarti, A. W. Chen, J. D. Varner, A review of the mammalian unfolded protein response. *Biotechnol. Bioeng.* **108**, 2777–2793 (2011).
19. E. Weber-ban, W. Kress, Clp chaperone e proteases : structure and function. **160**, 618–628 (2009).
20. S. M. Doyle, O. Genest, S. Wickner, Protein rescue from aggregates by powerful molecular chaperone machines. *Nat. Rev. Mol. Cell Biol.* **14**, 617–629 (2013).
21. D. Huber, B. Bukau, DegP: a Protein “Death Star.” *Structure.* **16**, 989–990 (2008).
22. O. Subrini, J. M. Betton, Assemblies of DegP underlie its dual chaperone and protease function. *FEMS Microbiol. Lett.* **296**, 143–148 (2009).
23. J. W. Fairman, N. Noinaj, S. K. Buchanan, The structural biology of β -barrel membrane proteins: A summary of recent reports. *Curr. Opin. Struct. Biol.* **21**, 523–531 (2011).
24. M. Riley, T. Abe, M. B. Arnaud, M. K. B. Berlyn, F. R. Blattner, R. R. Chaudhuri, J. D. Glasner, T. Horiuchi, I. M. Keseler, T. Kosuge, H. Mori, N. T. Perna, G. Plunkett, K. E. Rudd, M. H. Serres, G. H. Thomas, N. R. Thomson, D. Wishart, B. L. Wanner, Escherichia coli K-12: A cooperatively developed annotation snapshot - 2005. *Nucleic Acids Res.* **34**, 1–9 (2006).

25. C. Goemans, K. Denoncin, J. F. Collet, Folding mechanisms of periplasmic proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1517–1528 (2014).
26. H. Nikaido, M. Vaara, Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* **49**, 1–32 (1985).
27. T. Cranford-Smith, D. Huber, The way is the goal: How SecA transports proteins across the cytoplasmic membrane in bacteria. *FEMS Microbiol. Lett.* **365**, 1–16 (2018).
28. W. Wickner, The Enzymology Of Protein Translocation Across The Escherichia Coli Plasma Membrane. *Annu. Rev. Biochem.* **60**, 101–124 (1991).
29. M. Hiroyuki, I. Koreaki, The Sec protein-translocation pathway. *TRENDS Microbiol.* **9**, 494–499 (2001).
30. P. Natale, T. Brüser, A. J. M. Driessen, Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms. *Biochim. Biophys. Acta - Biomembr.* **1778**, 1735–1756 (2008).
31. G. von Heijne, The signal peptide. *J. Membr. Biol.* **115**, 195–201 (1990).
32. M. Pohlschröder, E. Hartmann, N. J. Hand, K. Dilks, A. Haddad, Diversity and evolution of protein translocation. *Annu. Rev. Microbiol.* **59**, 91–111 (2005).
33. H. Tokuda, S. Matsuyama, K. Tanaka-Masuda, in *The Periplasm* (ASM Press, Washington, DC, USA, 2014; <http://doi.wiley.com/10.1128/9781555815806.ch4>), pp. 67–79.
34. A. M. Plummer, K. G. Fleming, From Chaperones to the Membrane with a BAM ! *Trends Biochem. Sci.* **41**, 872–882 (2016).
35. A. E. Rizzitello, J. R. Harper, T. J. Silhavy, Genetic evidence for parallel pathways of chaperone activity in the periplasm of Escherichia coli. *J. Bacteriol.* **183**, 6794–6800 (2001).
36. M. P. Bos, V. Robert, J. Tommassen, Biogenesis of the gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* **61**, 191–214 (2007).
37. S. W. Lazar, R. Kolter, SurA assists the folding of Escherichia coli outer membrane proteins. *J. Bacteriol.* **178**, 1770–1773 (1996).

38. J. G. Sklar, T. Wu, D. Kahne, T. J. Silhavy, Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21**, 2473–2484 (2007).
39. J. Weski, M. Ehrmann, Genetic analysis of 15 protein folding factors and proteases of the Escherichia coli cell envelope. *J. Bacteriol.* **194**, 3225–3233 (2012).
40. A. N. Calabrese, B. Schiffrin, M. Watson, T. K. Karamanos, M. Walko, J. R. Humes, J. E. Horne, P. White, A. J. Wilson, A. C. Kalli, R. Tuma, A. E. Ashcroft, D. J. Brockwell, S. E. Radford, Inter-domain dynamics in the chaperone SurA and multi-site binding to its outer membrane protein clients. *Nat. Commun.* **11**, 1–16 (2020).
41. J. Thoma, B. M. Burmann, S. Hiller, D. J. Müller, Impact of holdase chaperones Skp and SurA on the folding of β -barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **22**, 795–802 (2015).
42. E. Braselmann, J. L. Chaney, M. M. Champion, P. L. Clark, DegP chaperone suppresses toxic inner membrane translocation intermediates. *PLoS One.* **11**, 1–22 (2016).
43. M. CastilloKeller, R. Misra, Protease-deficient DegP suppresses lethal effects of a mutant OmpC protein by its capture. *J. Bacteriol.* **185**, 148–154 (2003).
44. X. Ge, R. Wang, J. Ma, Y. Liu, A. N. Ezemaduka, P. R. Chen, X. Fu, Z. Chang, DegP primarily functions as a protease for the biogenesis of β -barrel outer membrane proteins in the Gram-negative bacterium Escherichia coli. *FEBS J.* **281**, 1226–1240 (2014).
45. G. Hansen, R. Hilgenfeld, Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response. *Cell. Mol. Life Sci.* **70**, 761–775 (2013).
46. E. B. Volokhina, J. Grijpstra, M. Stork, I. Schilders, J. Tommassen, M. P. Bos, Role of the periplasmic chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in Neisseria meningitidis. *J. Bacteriol.* **193**, 1612–1621 (2011).
47. B. Kern, O. P. Leiser, R. Misra, Suppressor Mutations in degS Overcome the Acute Temperature-Sensitive Phenotype of Δ degP and Δ degP Δ tol-pal Mutants of Escherichia coli. *J. Bacteriol.* **201**, 1–13 (2019).

48. T. A. Walton, C. M. Sandoval, C. A. Fowler, A. Pardi, M. C. Sousa, The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1772–1777 (2009).
49. B. M. Burmann, C. Wang, S. Hiller, Conformation and dynamics of the periplasmic membrane-protein–chaperone complexes OmpX–Skp and tOmpA–Skp. *Nat. Struct. Mol. Biol.* **20**, 1265–1272 (2013).
50. D. D. Isaac, J. S. Pinkner, S. J. Hultgren, T. J. Silhavy, The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. *Proc. Natl. Acad. Sci.* **102**, 17775–17779 (2005).
51. S. Quan, P. Koldewey, T. Tapley, N. Kirsch, K. M. Ruane, J. Pfizenmaier, R. Shi, S. Hofmann, L. Foit, G. Ren, U. Jakob, Z. Xu, M. Cygler, J. C. A. Bardwell, Articles Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nat. Publ. Gr.* **18**, 262–269 (2011).
52. F. A. Saul, J. P. Arié, B. Vulliez-le Normand, R. Kahn, J. M. Betton, G. A. Bentley, Structural and Functional Studies of FkpA from *Escherichia coli*, a cis/trans Peptidyl-prolyl Isomerase with Chaperone Activity. *J. Mol. Biol.* **335**, 595–608 (2004).
53. J. Schwalm, T. F. Mahoney, G. R. Soltes, T. J. Silhavy, Role for Skp in LptD assembly in *Escherichia coli*. *J. Bacteriol.* **195**, 3734–3742 (2013).
54. G. L. Thede, D. C. Arthur, R. A. Edwards, D. R. Buelow, J. L. Wong, T. L. Raivio, J. N. M. Glover, Structure of the periplasmic stress response protein CpxP. *J. Bacteriol.* **193**, 2149–2157 (2011).
55. K. S. Gajiwala, S. K. Burley, HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *J. Mol. Biol.* **295**, 605–612 (2000).
56. B. S. Mamathambika, J. C. Bardwell, Disulfide-linked protein folding pathways. *Annu. Rev. Cell Dev. Biol.* **24**, 211–235 (2008).
57. A. M. Mitchell, T. J. Silhavy, Envelope stress responses: balancing damage repair and toxicity. *Nat. Rev. Microbiol.* **17**, 417–428 (2019).
58. N. P. Walsh, B. M. Alba, B. Bose, C. A. Gross, R. T. Sauer, OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell.* **113**, 61–71 (2003).

59. Y. Akiyama, K. Kanehara, K. Ito, RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences. *EMBO J.* **23**, 4434–4442 (2004).
60. J. McEwen, P. Silverman, Chromosomal mutations of *Escherichia coli* that alter expression of conjugative plasmid functions. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 513–517 (1980).
61. D. Jianming, Iuchi Shiro, K. Hoi-Shan, L. Zhe, E. C. C. Lin, The deduced amino-acid sequence of the cloned cpxR gene suggests the protein is the cognate regulator for the membrane sensor, CpxA, in a two-component signal transduction system of *Escherichia coli*. *Gene*. **136**, 227–230 (1993).
62. S. L. Vogt, T. L. Raivio, Just scratching the surface: An expanding view of the Cpx envelope stress response. *FEMS Microbiol. Lett.* **326**, 2–11 (2012).
63. T. L. Raivio, T. J. Silhavy, The $\sigma(E)$ and Cpx regulatory pathways: Overlapping but distinct envelope stress responses. *Curr. Opin. Microbiol.* **2**, 159–165 (1999).
64. A. Delhay, J. F. Collet, G. Laloux, A Fly on the Wall: How Stress Response Systems Can Sense and Respond to Damage to Peptidoglycan. *Front. Cell. Infect. Microbiol.* **9** (2019), doi:10.3389/fcimb.2019.00380.
65. E. Di Cera, Serine proteases. *IUBMB Life*. **61**, 510–515 (2009).
66. L. Polgár, The catalytic triad of serine peptidases. *Cell. Mol. Life Sci.* **62** (2005), pp. 2161–2172.
67. L. Hedstrom, Serine protease mechanism and specificity. *Chem. Rev.* **102**, 4501–4523 (2002).
68. J. J. Perona, C. S. Craik, Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **4**, 337–360 (1995).
69. J. J. Perona, C. S. Craik, Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **272**, 29987–29990 (1997).
70. T. Clausen, C. Southan, M. Ehrmann, The HtrA family of proteases: Implications for protein composition and cell fate. *Mol. Cell.* **10**, 443–455 (2002).
71. M. B. Kennedy, Origin of PDZ (DHR, GLGF) domains. *Trends Biochem. Sci.* **20**, 350 (1995).

72. H. Schuhmann, P. F. Huesgen, I. Adamska, The family of Deg/HtrA proteases in plants. *BMC Plant Biol.* **12** (2012), doi:10.1186/1471-2229-12-52.
73. M. J. Pallen, B. W. Wren, The HtrA family of serine proteases. *Mol. Microbiol.* **26**, 209–221 (1997).
74. N. Singh, R. R. Kuppili, K. Bose, The structural basis of mode of activation and functional diversity: A case study with HtrA family of serine proteases. *Arch. Biochem. Biophys.* **516**, 85–96 (2011).
75. C. Eigenbrot, M. Ultsch, M. T. Lipari, P. Moran, S. J. Lin, R. Ganesan, C. Quan, J. Tom, W. Sandoval, M. Van Lookeren Campagne, D. Kirchhofer, Structural and functional analysis of HtrA1 and its subdomains. *Structure.* **20**, 1040–1050 (2012).
76. D. Zurawa-Janicka, J. Skorko-Glonek, B. Lipinska, HtrA proteins as targets in therapy of cancer and other diseases. *Expert Opin. Ther. Targets.* **14**, 665–679 (2010).
77. L. Truebestein, A. Tennstaedt, T. Mönig, T. Krojer, F. Canellas, M. Kaiser, T. Clausen, M. Ehrmann, Substrate-induced remodeling of the active site regulates human HTRA1 activity. *Nat. Struct. Mol. Biol.* **18**, 386–388 (2011).
78. Y. Suzuki, Y. Imai, H. Nakayama, K. Takahashi, K. Takio, R. Takahashi, A serine protease, HtrA2, is released from the mitochondria and interacts with XIAP, inducing cell death. *Mol. Cell.* **8**, 613–621 (2001).
79. L. Vande Walle, M. Lamkanfi, P. Vandenabeele, The mitochondrial serine protease HtrA2/Omi: an overview. *Cell Death Differ.* **15**, 453–460 (2008).
80. W. Li, S. M. Srinivasula, J. Chai, P. Li, J. W. Wu, Z. Zhang, E. S. Alnemri, Y. Shi, Structural insights into the pro-apoptotic function of mitochondrial serine protease htra2/omi. *Nat. Struct. Biol.* **9**, 436–441 (2002).
81. G. Y. Nie, A. Hampton, Y. Li, J. K. Findlay, L. A. Salamonsen, Identification and cloning of two isoforms of human high-temperature requirement factor A3 (HtrA3), characterization of its genomic structure and comparison of its tissue distribution with HtrA1 and HtrA2. *Biochem. J.* **371**, 39–48 (2003).
82. P. Glaza, J. Osipiuk, T. Wenta, D. Zurawa-Janicka, M. Jarzab, A. Lesner, B. Banecki, J. Skorko-Glonek, A. Joachimiak, B. Lipinska, Structural and functional analysis of human HtrA3 protease and its subdomains. *PLoS One.* **10**, 1–24 (2015).

83. T. Wenta, M. Jarzab, M. Rychlowski, M. Borysiak, A. Latala, D. Zurawa-Janicka, A. Filipek, B. Lipinska, Cellular substrates and pro-apoptotic function of the human HtrA4 protease. *J. Proteomics*. **209**, 103505 (2019).
84. H. Schuhmann, I. Adamska, Deg proteases and their role in protein quality control and processing in different subcellular compartments of the plant cell. *Physiol. Plant*. **145**, 224–234 (2012).
85. J. Kley, B. Schmidt, B. Boyanov, P. C. Stolt-Bergner, R. Kirk, M. Ehrmann, R. R. Knopf, L. Naveh, Z. Adam, T. Clausen, Structural adaptation of the plant protease Deg1 to repair photosystem II during light exposure. *Nat. Struct. Mol. Biol.* **18**, 728–731 (2011).
86. R. Sun, H. Fan, F. Gao, Y. Lin, L. Zhang, W. Gong, L. Liu, Crystal structure of Arabidopsis Deg2 protein reveals an internal PDZ ligand locking the hexameric resting state. *J. Biol. Chem.* **287**, 37564–37569 (2012).
87. M. Ouyang, X. Li, S. Zhao, H. Pu, J. Shen, Z. Adam, T. Clausen, L. Zhang, The crystal structure of Deg9 reveals a novel octameric-type HtrA protease. *Nat. Plants*. **3**, 973–982 (2017).
88. O. Cheregi, R. Wagner, C. Funk, Insights into the Cyanobacterial Deg/HtrA Proteases. *Front. Plant Sci.* **7**, 694 (2016).
89. H. Kolmar, P. R. H. Waller, R. T. Sauer, The DegP and DegQ periplasmic endoproteases of Escherichia coli: Specificity for cleavage sites and substrate conformation. *J. Bacteriol.* **178**, 5925–5929 (1996).
90. T. Clausen, M. Kaiser, R. Huber, M. Ehrmann, HTRA proteases: regulated proteolysis in protein quality control. *Nat. Rev. Mol. Cell Biol.* **12**, 152–162 (2011).
91. S. S. Kim, I. Song, G. Eom, S. S. Kim, A small periplasmic protein with a hydrophobic C-terminal residue enhances DegP proteolysis as a suicide activator. *J. Bacteriol.* **200**, 1–12 (2018).
92. J. Skórko-Glonek, a Wawrzynow, K. Krzewski, K. Kurpierz, B. Lipinska, A. Wawrzynow, Site-directed mutagenesis of the HtrA (DegP) serine-protease, whose proteolytic activity is indispensable for Escherichia coli survival at elevated-temperatures. *Gene*. **163**, 47–52 (1995).

93. H. Malet, F. Canellas, J. Sawa, J. Yan, K. Thalassinou, M. Ehrmann, T. Clausen, H. R. Saibil, Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ. *Nat. Struct. Mol. Biol.* **19**, 152–157 (2012).
94. K. Fang, X. Jin, S. H. Hong, Probiotic *Escherichia coli* inhibits biofilm formation of pathogenic *E. coli* via extracellular activity of DegP. *Sci. Rep.* **8**, 4939 (2018).
95. C. M. Abfalter, M. Schubert, C. Götz, T. P. Schmidt, G. Posselt, S. Wessler, HtrA-mediated E-cadherin cleavage is limited to DegP and DegQ homologs expressed by gram-negative pathogens. *Cell Commun. Signal.* **14**, 30 (2016).
96. S. Backert, S. Bernegger, J. Skórko-Glonek, S. Wessler, Extracellular HtrA serine proteases: An emerging new strategy in bacterial pathogenesis. *Cell. Microbiol.* **20**, 1–9 (2018).
97. T. Krojer, M. Garrido-Franco, R. Huber, M. Ehrmann, T. Clausen, Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature.* **416**, 455–459 (2002).
98. T. Krojer, J. Sawa, R. Huber, T. Clausen, HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. *Nat. Struct. Mol. Biol.* **17**, 844–852 (2010).
99. T. Krojer, K. Pangerl, J. Kurt, J. Sawa, C. Stingl, K. Mechtler, R. Huber, M. Ehrmann, T. Clausen, Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins. *Proc. Natl. Acad. Sci.* **105**, 7702–7707 (2008).
100. S. Kim, R. A. Grant, R. T. Sauer, Covalent linkage of distinct substrate degrons controls assembly and disassembly of DegP proteolytic cages. *Cell.* **145**, 67–78 (2011).
101. A. Sobiecka-Szkatula, A. Polit, A. Scire, A. Gieldon, F. Tanfani, Z. Szkarlat, J. Ciarkowski, D. Zurawa-Janicka, J. Skorko-Glonek, B. Lipinska, Temperature-induced conformational changes within the regulatory loops L1-L2-LA of the HtrA heat-shock protease from *Escherichia coli*. *Biochim. Biophys. Acta - Proteins Proteomics.* **1794**, 1573–1582 (2009).
102. A. Sobiecka-Szkatula, A. Gieldon, A. Scire, F. Tanfani, D. Figaj, T. Koper, J. Ciarkowski, B. Lipinska, J. Skorko-Glonek, The role of the L2 loop in the regulation

- and maintaining the proteolytic activity of HtrA (DegP) protein from *Escherichia coli*. *Arch. Biochem. Biophys.* **500**, 123–130 (2010).
103. D. Figaj, A. Gieldon, M. Bartczak, T. Koper, U. Zarzecka, A. Lesner, B. Lipinska, J. Skorko-Glonek, The LD loop as an important structural element required for transmission of the allosteric signal in the HtrA (DegP) protease from *Escherichia coli*. *FEBS J.* **283**, 3471–3487 (2016).
 104. T. Krojer, J. Sawa, E. Schäfer, H. R. Saibil, M. Ehrmann, T. Clausen, Structural basis for the regulated protease and chaperone function of DegP. *Nature.* **453**, 885–890 (2008).
 105. S. Kim, R. T. Sauer, Cage assembly of DegP protease is not required for substrate-dependent regulation of proteolytic activity or high-temperature cell survival. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 7263–7268 (2012).
 106. N. J. Thompson, M. Merdanovic, M. Ehrmann, E. Van Duijn, A. J. R. Heck, E. Van Duijn, A. J. R. Heck, Substrate occupancy at the onset of oligomeric transitions of DegP. *Structure.* **22**, 281–290 (2014).
 107. B. M. Alba, C. A. Gross, Regulation of the *Escherichia coli* σ E-dependent envelope stress response. *Mol. Microbiol.* **52**, 613–619 (2004).
 108. M. Meltzer, S. Hasenbein, P. Hauske, N. Kucz, M. Merdanovic, S. Grau, A. Beil, D. Jones, T. Krojer, T. Clausen, M. Ehrmann, M. Kaiser, Allosteric activation of HtrA protease DegP by stress signals during bacterial protein quality control. *Angew. Chemie - Int. Ed.* **47**, 1332–1334 (2008).
 109. J. Sambrook, E. F. Fritsch, T. Maniatis, Molecular cloning: a laboratory manual. *Mol. cloning a Lab. manual.* (1989).
 110. L. Zheng, U. Baumann, J. L. Reymond, An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32** (2004), doi:10.1093/nar/gnh110.
 111. M. P. Weiner, G. L. Costa, W. Schoettlin, J. Cline, E. Mathur, J. C. Bauer, Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene.* **151**, 119–123 (1994).
 112. P. Starokadomskyy, E. Burstein, BAP: a variation of TAP with Multiple Application.

- Methods Mol. Biol.* **1177**, 193–209 (2014).
113. F. W. Studier, B. A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130 (1986).
 114. C. G. Kalodimos, Structure and Flexibility Adaptation in Nonspecific and Specific Protein-DNA Complexes. *Science (80-.)*. **305**, 386–389 (2004).
 115. F. Baneyx, Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10**, 411–421 (1999).
 116. J. H. Miller, Experiments in molecular genetics (1972).
 117. M. Sattler, S. W. Fesik, Use of deuterium labeling in NMR: Overcoming a sizeable problem. *Structure*. **4**, 1245–1249 (1996).
 118. V. Tugarinov, L. E. Kay, An Isotope Labeling Strategy for Methyl TROSY Spectroscopy. *J. Biomol. NMR*. **28**, 165–172 (2004).
 119. I. Ayala, R. Sounier, N. Usé, P. Gans, J. Boisbouvier, An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J. Biomol. NMR*. **43**, 111–119 (2009).
 120. I. Gelis, A. M. J. J. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou, C. G. Kalodimos, Structural Basis for Signal-Sequence Recognition by the Translocase Motor SecA as Determined by NMR. *Cell*. **131**, 756–769 (2007).
 121. R. Sprangers, L. E. Kay, Probing supramolecular structure from measurement of methyl ¹H-¹³C residual dipolar couplings. *J. Am. Chem. Soc.* **129**, 12668–12669 (2007).
 122. R. Kerfah, M. J. Plevin, R. Sounier, P. Gans, J. Boisbouvier, Methyl-specific isotopic labeling: A molecular tool box for solution NMR studies of large proteins. *Curr. Opin. Struct. Biol.* **32**, 113–122 (2015).
 123. J. Porath, J. Carlsson, I. Olsson, G. Belfrage, Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*. **258**, 598–599 (1975).
 124. H. Block, B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, F. Schäfer, Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). A Review. *Methods Enzymol.* **463**, 439–473 (2009).

125. M. E. Kimple, A. L. Brill, R. L. Pasker, Overview of affinity tags for protein purification. *Curr. Protoc. Protein Sci.*, 1–23 (2013).
126. J.-C. Janson, *Protein Purification* (John Wiley & Sons, Inc., Hoboken, NJ, USA, ed. 3, 2011; <http://doi.wiley.com/10.1002/9780470939932>), *Methods of Biochemical Analysis*.
127. J. Porath, P. Flodin, Gel Filtration: A Method for Desalting and Group Separation. *Nature*. **183**, 1657–1659 (1959).
128. V. Tugarinov, V. Kanelis, L. E. Kay, Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. **1**, 749–754 (2006).
129. M. Li, Z.-G. Su, J.-C. Janson, In vitro protein refolding by chromatographic procedures. *Protein Expr. Purif.* **33**, 1–10 (2004).
130. A. P. J. Middelberg, Preparative protein refolding. *Trends Biotechnol.* **20**, 437–43 (2002).
131. H. Lilie, E. Schwarz, R. Rudolph, Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.* **9**, 497–501 (1998).
132. T. J. Beveridge, Structures of gram-negative cell walls and their derived membrane vesicles. *J. Bacteriol.* **181**, 4725–4733 (1999).
133. K. E. Bonnington, M. J. Kuehn, Protein selection and export via outer membrane vesicles. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1612–1619 (2014).
134. N. C. Kesty, M. J. Kuehn, Incorporation of Heterologous Outer Membrane and Periplasmic Proteins into *Escherichia coli* Outer Membrane Vesicles. *J. Biol. Chem.* **279**, 2069–2076 (2004).
135. J. Thoma, S. Manioglou, D. Kalbermatter, P. D. Bosshart, D. Fotiadis, D. J. Müller, Protein-enriched outer membrane vesicles as a native platform for outer membrane protein studies. *Commun. Biol.* **1** (2018), doi:10.1038/s42003-018-0027-5.
136. J. Thoma, B. M. Burmann, High-Resolution In Situ NMR Spectroscopy of Bacterial Envelope Proteins in Outer Membrane Vesicles. *Biochemistry*. **59**, 1656–1660 (2020).
137. H. M. Schmitt, A. Brecht, J. Piehler, G. Gauglitz, An integrated system for optical biomolecular interaction analysis. *Biosens. Bioelectron.* **12**, 809–816 (1997).

138. J. Concepcion, K. Witte, C. Wartchow, S. Choo, D. Yao, H. Persson, J. Wei, P. Li, B. Heidecker, W. Ma, R. Varma, L.-S. Zhao, D. Perillat, G. Carricato, M. Recknor, K. Du, H. Ho, T. Ellis, J. Gamez, M. Howes, J. Phi-Wilson, S. Lockard, R. Zuk, H. Tan, Label-free detection of biomolecular interactions using biolayer interferometry for kinetic characterization. *Comb. Chem. High Throughput Screen.* **12**, 791–800 (2009).
139. S. Patching, NMR-Active Nuclei for Biological and Biomedical Applications. *J. Diagnostic Imaging Ther.* **3**, 7–48 (2016).
140. M. H. Levitt, *Spin dynamics: basics of nuclear magnetic resonance* (John Wiley & Sons, 2013).
141. J. Clayden, N. Greeves, S. G. Warren, *Organic chemistry Jonathan Clayden, Nick Greeves, Stuart Warren.* (Oxford University Press, Oxford ; New York, Second edi., 2012).
142. G. S. Rule, T. K. Hitchens, *Fundamentals of protein NMR spectroscopy* (Springer Science & Business Media, 2006), vol. 5.
143. A. D. Kline, W. Braun, K. Wüthrich, Determination of the complete three-dimensional structure of the α -amylase inhibitor tendamistat in aqueous solution by nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **204**, 675–724 (1988).
144. T. D. W. Claridge, Chapter 8 Correlations through space. The nuclear Overhauser effect. *Tetrahedron Org. Chem. Ser.* **27**, 247–302 (2009).
145. J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553 (1979).
146. D. S. Wishart, B. D. Sykes, F. M. Richards, The Chemical Shift Index : A Fast and Simple Method for the Assignment of Protein Secondary Structure Through NMR Spectroscopy. *Biochemistry.* **31**, 1647–1651 (1992).
147. D. S. Wishart, B. D. Sykes, The ^{13}C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR.* **4**, 171–180 (1994).
148. J. T. Nielsen, F. A. A. Mulder, Potenci: Prediction of temperature, neighbor and ph-corrected chemical shifts for intrinsically disordered proteins. *J. Biomol. NMR.* **70**, 141–165 (2018).

149. P. Güntert, Automated structure determination from NMR spectra. *Eur. Biophys. J.* **38**, 129–143 (2009).
150. K. Kazimierczuk, V. Orekhov, Non-uniform sampling: Post-Fourier era of NMR data collection and processing. *Magn. Reson. Chem.* **53**, 921–926 (2015).
151. K. Kazimierczuk, V. Y. Orekhov, Accelerated NMR spectroscopy by using compressed sensing. *Angew. Chemie - Int. Ed.* **50**, 5556–5559 (2011).
152. S. Bagby, Encyclopedia of Biophysics. *Encycl. Biophys.* (2013), doi:10.1007/978-3-642-16712-6.
153. E. Schmidt, J. Gath, B. Habenstein, F. Ravotti, K. Székely, M. Huber, L. Buchner, A. Böckmann, B. H. Meier, P. Güntert, Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. *J. Biomol. NMR.* **56**, 243–254 (2013).
154. I. Pritišanac, J. M. Würz, T. R. Alderson, P. Güntert, Automatic structure-based NMR methyl resonance assignment in large proteins. *Nat. Commun.* **10**, 1–12 (2019).
155. P. Güntert, Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378 (2004).
156. K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins. *Nature.* **450**, 964–972 (2007).
157. H. Frauenfelder, S. G. Sligar, P. G. Wolynes, The Energy Landscapes and Motions of Proteins. *Science (80-.).* **254**, 1598–1603 (1991).
158. L. E. Kay, Protein dynamics from NMR. *Biochem. Cell Biol.* **76**, 145–152 (1998).
159. A. G. Palmer, NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004).
160. H. Saitô, I. Ando, A. Ramamoorthy, Chemical shift tensor - The heart of NMR: Insights into biological aspects of proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **57**, 181–228 (2010).
161. V. Kharchenko, M. Nowakowski, M. Jaremko, A. Ejchart, Ł. Jaremko, Dynamic $^{15}\text{N}\{^1\text{H}\}$ NOE measurements: a tool for studying protein dynamics. *J. Biomol. NMR.* **74**, 707–716 (2020).

162. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, K. Wuthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366–12371 (1997).
163. N. Salvi, *Dynamic Studies Through Control of Relaxation in NMR Spectroscopy* (2014; <http://link.springer.com/10.1007/978-3-319-06170-2>).
164. D. M. Korzhnev, K. Klover, V. Kanelis, V. Tugarinov, L. E. Kay, Probing Slow Dynamics in High Molecular Weight Proteins by Methyl-TROSY NMR Spectroscopy: Application to a 723-Residue Enzyme. *J. Am. Chem. Soc.* **126**, 3964–3973 (2004).
165. T. R. Alderson, L. E. Kay, NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell*. **184**, 577–595 (2021).
166. G. Lipari, A. Szabo, Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **104**, 4546–4559 (1982).
167. G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, A. M. Gronenborn, Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. *J. Am. Chem. Soc.* **112**, 4989–4991 (1990).
168. E. J. d’Auvergne, P. R. Gooley, The use of model selection in the model-free analysis of protein dynamics. *J. Biomol. NMR.* **25**, 25–39 (2003).
169. P. Dosset, J. C. Hus, M. Blackledge, D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR.* **16**, 23–28 (2000).
170. R. Cole, J. P. Loria, FAST-Modelfree: A program for rapid automated analysis of solution NMR spin-relaxation data. *J. Biomol. NMR.* **26**, 203–213 (2003).
171. S. Morin, T. E. Linnet, M. Lescanne, P. Schanda, G. S. Thompson, M. Tollinger, K. Teilmann, S. Gagné, D. Marion, C. Griesinger, M. Blackledge, E. J. D’auvergne, Relax: The analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data. *Bioinformatics.* **30**, 2219–2220 (2014).
172. M. J. Hartl, B. M. Burmann, S. J. Prasad, C. Schwarzing, K. Schweimer, B. M.

- Wöhrl, P. Rösch, S. Schwarzing, Fast mapping of biomolecular interfaces by random spin labeling (rsl). *J. Biomol. Struct. Dyn.* **29**, 793–798 (2012).
173. J. L. Battiste, G. Wagner, Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry*. **39**, 5355–5365 (2000).
 174. N. U. Jain, A. Venot, K. Umemoto, H. Leffler, J. H. Prestegard, Distance mapping of protein-binding sites using spin-labeled oligosaccharide ligands. *Protein Sci.* **10**, 2393–2400 (2008).
 175. M. Sjødt, R. Clubb, Nitroxide Labeling of Proteins and the Determination of Paramagnetic Relaxation Derived Distance Restraints for NMR Studies. *BIO-PROTOCOL*. **7** (2017), doi:10.21769/BioProtoc.2207.
 176. E. R. P. Zuiderweg, Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry*. **41**, 1–7 (2002).
 177. F. H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K. P. Neidig, H. R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR*. **39**, 275–289 (2007).
 178. H. Matsuo, K. J. Walters, K. Teruya, T. Tanaka, G. T. Gassner, S. J. Lippard, Y. Kyogoku, G. Wagner, Identification by NMR spectroscopy of residues at contact surfaces in large, slowly exchanging macromolecular complexes [17]. *J. Am. Chem. Soc.* **121**, 9903–9904 (1999).
 179. E. O. Stejskal, J. E. Tanner, Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* **42**, 288–292 (1965).
 180. A. J. Simpson, Determining the molecular weight, aggregation, structures and interactions of natural organic matter using diffusion ordered spectroscopy. *Magn. Reson. Chem.* **40**, 72–82 (2002).
 181. I. Keresztes, P. G. Williard, V. Pro, R. Island, R. V June, Diffusion-Ordered NMR Spectroscopy (DOSY) of THF Solvated n -Butyllithium Aggregates Organolithium reagents have been used extensively in synthetic organic chemistry . Considerable amount of work has been done on the structural characterization of organ, 10228–

- 10229 (2000).
182. C. S. Johnson, Diffusion ordered nuclear magnetic resonance spectroscopy: Principles and applications. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 203–256 (1999).
 183. J. J. Chou, J. L. Baber, A. Bax, Characterization of phospholipid mixed micelles by translational diffusion. *J. Biomol. NMR.* **29**, 299–308 (2004).
 184. D. Apperley, R. Harris, P. Hodgkinson, *Solid-State NMR: Basic Principles & Practice* (Momentum Press, Oxford, UK, 2012; <http://doi.wiley.com/10.1002/9780470999394>).
 185. J. Herzfeld, A. E. Berger, Sideband intensities in NMR spectra of samples spinning at the magic angle. *J. Chem. Phys.* **73**, 6021–6030 (1980).
 186. P. C. A. Van Der Wel, New applications of solid-state NMR in structural biology. *Emerg. Top. Life Sci.* **2**, 57–67 (2018).
 187. V. A. Higman, Solid-state MAS NMR resonance assignment methods for proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **106–107**, 37–65 (2018).
 188. P. Schanda, M. Ernst, Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **96**, 1–46 (2016).
 189. B. Reif, S. E. Ashbrook, L. Emsley, M. Hong, Solid-state NMR spectroscopy. *Nat. Rev. Methods Prim.* **1** (2021), doi:10.1038/s43586-020-00002-1.
 190. J. Jiang, K. Prasad, E. M. Lafer, R. Sousa, Structural basis of interdomain communication in the Hsc70 chaperone. *Mol. Cell.* **20**, 513–524 (2005).
 191. M. P. Foster, C. A. Mcelroy, C. D. Amero, Current Topics Solution NMR of Large Molecules and Assemblies †. *Current.* **46** (2007).
 192. J. G. De La Torre, M. L. Huertas, B. Carrasco, HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *J. Magn. Reson.* **147**, 138–146 (2000).
 193. C. C. Valley, A. Cembran, J. D. Perlmutter, A. K. Lewis, N. P. Labello, J. Gao, J. N. Sachs, The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J. Biol. Chem.* **287**, 34979–34991 (2012).

194. M. J. Plevin, D. L. Bryce, J. Boisbouvier, Direct detection of CH/ π interactions in proteins. *Nat. Chem.* **2**, 466–471 (2010).
195. R. W. Harkness, Y. Toyama, Z. A. Ripstein, H. Zhao, A. I. M. Sever, Q. Luan, J. P. Brady, P. L. Clark, P. Schuck, L. E. Kay, Competing stress-dependent oligomerization pathways regulate self-assembly of the periplasmic protease-chaperone DegP. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021), doi:10.1073/pnas.2109732118.
196. D. Figaj, A. Gieldon, A. Polit, A. Sobiecka-Szkatula, T. Koper, M. Denkiewicz, B. Banecki, A. Lesner, J. Ciarkowski, B. Lipinska, J. Skorko-Glonek, The la loop as an important regulatory element of the HtrA (DegP) protease from Escherichia coli structural and functional studies. *J. Biol. Chem.* **289**, 15880–15893 (2014).
197. A. K. De Regt, S. Kim, J. Sohn, R. A. Grant, T. A. Baker, R. T. Sauer, A conserved activation cluster is required for allosteric communication in htra-family proteases. *Structure.* **23**, 517–526 (2015).
198. J. P. Demers, P. Fricke, C. Shi, V. Chevelkov, A. Lange, Structure determination of supra-molecular assemblies by solid-state NMR: Practical considerations. *Prog. Nucl. Magn. Reson. Spectrosc.* **109**, 51–78 (2018).
199. C. Dartigalongue, D. Missiakas, S. Raina, Characterization of the Escherichia coli σ E Regulon. *J. Biol. Chem.* **276**, 20866–20875 (2001).
200. T. L. Raivio, S. K. D. Leblanc, N. L. Price, The Escherichia coli Cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. *J. Bacteriol.* **195**, 2755–2767 (2013).
201. S. Wu, X. Ge, Z. Lv, Z. Zhi, Z. Chang, X. S. Zhao, Interaction between bacterial outer membrane proteins and periplasmic quality control factors: A kinetic partitioning mechanism. *Biochem. J.* **438**, 505–511 (2011).
202. D. M. MacRitchie, N. Acosta, T. L. Raivio, DegP is involved in Cpx-mediated posttranscriptional regulation of the type III secretion apparatus in enteropathogenic Escherichia coli. *Infect. Immun.* **80**, 1766–1772 (2012).
203. T. Krojer, M. Garrido-franco, R. Huber, M. Ehrmann, T. Clausen, Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. **530**, 527–530 (2002).
204. Y. Toyama, R. W. Harkness, T. Y. T. Lee, J. T. Maynes, L. E. Kay, Oligomeric

assembly regulating mitochondrial HtrA2 function as examined by methyl-TROSY NMR. *Proc. Natl. Acad. Sci.* **118**, e2025022118 (2021).